DIRICHLET'S UNIT THEOREM

BEN MACKAY

Abstract. These notes aim to give a comprehensive review of the preliminaries needed to understand and prove the following version of Dirichlet's Unit Theorem:

$$\mathcal{O}_L \cong \mu_L \times \mathbb{Z}^{r+s-1},$$

for a number field L with $r+2s$ complex embeddings, where μ_L is the group of roots of unity of L, and $\mathcal{O}_L$ is the ring of algebraic integers of L. Furthermore, these notes are based on Jack Thorne's lecture notes for the 2019 delivery of the Number Fields course at Cambridge University, which in turn often reference Marcus' book *Number Fields*.

1. Field Extensions and Polynomial Rings

Definition 1.1. A *field extension* is an inclusion of fields $K \subseteq L$. That is, a pair of fields K, L such that $K \subseteq L$ and K is a field such that the operations of K are those of L, restricted to K. The notation L/K is often used to refer to a field extension.

Definition 1.2. Let L/K be a field extension. Note that L is a vector space over K. The *degree* of the field extension is

$$[L : K] = \dim_K(L).$$

Remark 1.3. It is important to note that the degree of a field extension is always non-zero. By definition, any field M has at least one non-zero element (the multiplicative identity 1). Let M/K be a field extension. Recall that $[M : K]$ is the dimension of M as a vector space over K. The only vector space of dimension zero is $\{0\}$, and we have just explained that M contains a non-zero element. Hence $[M : K] \geq 1$ for any field extension M/K.

Lemma 1.4. *Let M/K be a field extension. Then $[M : K] = 1 \iff M = K$.*

Proof. Suppose $M = K$. Then $\{1\}$ is a basis of M as a vector space over K, as each element of M can be written as $1 \cdot k$ for some $k \in K$. Hence $[M : K] = 1$. Conversely, if $[M : K] = 1$, then $\{1\}$ is a basis of M as a vector space over K: it is a linearly independent set, as

$$k \cdot 1 = 0 \implies k = 0,$$

since K is a field ($1 \neq 0$ and there are no zero-divisors), and it has $1 = \dim_K(M)$ elements. Thus, every element of M can be written as $k \cdot 1 = k$ for some $k \in K$, and thus $M = K$. □

Definition 1.5. A field extension with finite degree is called a *finite field extension.*

Date: February 21, 2026.

Theorem 1.6 (Tower Law). *Given a tower $K \hookrightarrow L \hookrightarrow M$ of field extensions,*

$$[M : K] = [M : L][L : K].$$

Proof. Let $(u_i)_{i \in I}$ be a basis for M over L and let $(v_j)_{j \in J}$ be a basis for L over K. Let $x \in M$ be a vector. Then we can write

$$x = \sum_{i \in I} \mu_i u_i$$

for some collection $(\mu_i)_{i \in I}$ of elements of L. Now, since $(v_j)_{j \in J}$ is a basis for L over K, we can write each μ_i as a linear combination of the elements of this basis. In other words, for each $i \in I$, we can write

$$\mu_i = \sum_{j \in J} \lambda_{ij} v_j$$

for some collection $(\lambda_{ij})_{j \in J}$ of elements of K. Thus, we can write

$$x = \sum_{i \in I} \sum_{j \in J} \lambda_{ij} u_i v_j.$$

Since $x \in M$ was taken to be arbitrary, it follows that $(u_i v_j)_{i \in I, j \in J}$ spans M over K (recall that this makes sense, as $u_i \in M$, $v_j \in L \subseteq M$, so $u_i v_j \in M$ for each $i \in I$ and $j \in J$, as M is a field and is hence closed under multiplication). Now, suppose that

$$\sum_{i \in I} \sum_{j \in J} \lambda_{ij} u_i v_j = 0$$

for some $(\lambda_{ij})_{i \in I, j \in J} \in K$. Then

$$\sum_{i \in I} \left(\sum_{j \in J} \lambda_{ij} v_j \right) u_i = 0,$$

and since $(u_i)_{i \in I}$ is a basis for M over L (and thus a linearly independent set of vectors of M over L), we must have that

$$\sum_{j \in J} \lambda_{ij} v_j = 0$$

for each $j \in J$. Now, since $(v_j)_{j \in J}$ is a basis for L over K, we must have that $(\lambda_{ij}) = 0$ for each $i \in I, j \in J$. Note that we have just shown that $(u_i v_j)_{i \in I, j \in J}$ is a linearly independent, spanning set of vectors of M over K, and is thus a basis of M over K. Thus,

$$[M : K] = \dim_K(M) = |(u_i v_j)_{i \in I, j \in J}| = |I| \cdot |J| = [M : L][L : K],$$

as required. □

Definition 1.7.

(1) Let L/K be a field extension. An element $\alpha \in L$ is said to be *algebraic over K* if there exists a monic polynomial $f \in K[x]$ such that $f(\alpha) = 0$.

(2) Let $L/\mathbb{Q}$ be a field extension. An *algebraic integer* is an element $\alpha \in L$ such that $f(\alpha) = 0$ for some monic polynomial $f \in \mathbb{Z}[x]$. The set of algebraic integers of L is denoted $\mathcal{O}_L$.

Definition 1.8. Let L/K be a field extension and $\alpha \in L$ be algebraic over K. The minimal polynomial f_α of α is the monic polynomial $f_\alpha \in K[x]$ of least degree such that $f_\alpha(\alpha) = 0$.

Proposition 1.9 (Euclidean algorithm for polynomials). *Let K be a field, and $f, g \in K[x]$. Then there exist $r, q \in K[x]$ such that*

$$f = gq + r,$$

with $\deg r < \deg g$.

Proof. Let $\deg f = n$, $\deg g = m$, and write

$$f(x) = \sum_{i=0}^{n} a_i x^i, \quad g(x) = \sum_{i=0}^{m} b_i x_i,$$

for some $a_0, \dots, a_n, b_0, \dots, b_m \in K$ with $a_n, b_m \neq 0$. If $n < m$, we let $q = 0$ and $r = f$, and are done. Otherwise, suppose $n \geq m$. We proceed by induction. Let

$$f_1 = f - a_n b_m^{-1} x^{n-m} g.$$

Then $f_1 \in K[x]$, as $K[x]$ is a ring and each element of K has a multiplicative inverse (as K is a field). Furthermore, the coefficient of x^n of f_1 is

$$a_n x^n - a_n b_m^{-1} x^{n-m} b_m x^m = a_n x^n - a_n b_m^{-1} b_m x^n = 0,$$

so $\deg f_1 < n$. Now, if $n = m$, then $\deg f_1 < n = m$, and

$$f = g(a_n b_m^{-1} x^{n-m}) + f_1,$$

and $\deg f_1 < \deg f$, so the base case holds. Now, let $n > m$ and suppose the statement holds for all $k < n$. Then, as $\deg f_1 < n$, we can write

$$f_1 = gq_1 + r_1,$$

for some $r_1, q_1 \in K[x]$ with $\deg r_1 < \deg g = m$. Thus,

$$f = g(a_n b_m^{-1} x^{n-m}) + gq_1 + r_1 = g(a_n b_m^{-1} x^{n-m} + q_1) + r_1,$$

and the statement holds for all $f, g \in K[x]$ by induction. □

Definition 1.10. Let $f \in \mathbb{Z}[x]$, and write

$$f(x) = a_n x^n + \cdots + a_1 x + a_0,$$

for some $n \in \mathbb{N}$ and $a_0, \dots, a_n \in \mathbb{Z}$. The *content* $c(f)$ of f is defined by

$$c(f) = \gcd(a_0, \dots, a_n).$$

Lemma 1.11 (Gauss' Lemma). *Let $f, g \in \mathbb{Z}[x]$. Then $c(fg) = c(f)c(g)$.*

Proof.

□

Lemma 1.12. *Let $L/\mathbb{Q}$ be a field extension, and let $\alpha \in L$ be an algebraic integer.*

(1) The minimal polynomial f_α of α over $\mathbb{Q}$ is contained in $\mathbb{Z}[x]$.

(2) If $g \in \mathbb{Z}[x]$ is any polynomial such that $g(\alpha) = 0$, then we can find $q \in \mathbb{Z}[x]$ such that $g = qf_\alpha$.

Proof. (1) Let $f \in \mathbb{Z}[x]$ be a monic polynomial such that $f(\alpha) = 0$. Then $f \in \mathbb{Q}[x]$ and hence (by the Euclidean algorithm for polynomials) we can find $q, r \in \mathbb{Q}[x]$ such that $f = qf_\alpha + r$, with $\deg r < \deg f_\alpha$. Note then that

$$f(\alpha) = q(\alpha)f_\alpha(\alpha) + r(\alpha) = 0 \implies r(\alpha) = 0 \implies r = 0,$$

as this would otherwise contradict the minimality of f_α. Let n, m be positive integers such that $nq, mf_\alpha \in \mathbb{Z}[x]$ (these exist, as we could take them to be the respective lowest common multiples of the denominators of the rational coefficients of each polynomial). By Gauss' Lemma, we have $nm = c(nmf) = c(nqmf_\alpha) = c(nq)c(mf_\alpha)$ (as f is monic, so $c(f) = 1$ as the leading coefficient of f is 1). Since f and f_α are monic, $f = qf_\alpha \implies q$ is monic. Hence, $c(nq) = c(n)c(q) = |n|$, $c(mf_\alpha) = c(m)c(f_\alpha) = c(m) = |m|$. Thus, $c(nq) \mid n$, $c(mf_\alpha) \mid m$. As $c(nq)c(mf_\alpha) = nm$, we must then have $c(nq) = n$, $c(mf_\alpha) = m$. Thus, we have

$$f_\alpha = \frac{1}{m}(mf_\alpha) \in \mathbb{Z}[x],$$

as dividing each coefficient of mf_α by $c(mf_\alpha)$ must give a polynomial with integer coefficients (each coefficent of mf_α is divisible by $c(mf_\alpha)$).

(2) Let $g \in \mathbb{Z}[x]$ be a non-zero polynomial such that $g(\alpha) = 0$ (if $g = 0$, we can take $q = 0$ and the proof is trivial). Then we can write $g = qf_\alpha + r$ for some $q, r \in \mathbb{Q}[x]$, with $\deg r < \deg f_\alpha$. By the same logic as before, we must have $r = 0$. Let $n \geq 1$ be an integer such that $nq \in \mathbb{Z}[x]$. Then $c(ng) = nc(g) = c(nqf_\alpha) = c(nq)$. Thus $n \mid c(nq)$, and we must have $q \in \mathbb{Z}[x]$. □

Corollary 1.13. $\mathcal{O}_\mathbb{Q} = \mathbb{Z}$.

Proof. If $\alpha \in \mathbb{Q}$, its minimal polynomial is $f_\alpha \in \mathbb{Q}[x]$, defined by $f_\alpha(x) = x - \alpha$. By the above lemma, we must have $\alpha \in \mathbb{Z}$. □

Proposition 1.14. *Let $L/\mathbb{Q}$ be a field extension. Then $\mathcal{O}_L$ is a ring.*

Proof. Clearly, $0, 1 \in \mathcal{O}_L$, as

$$f(0) = 0, \quad g(1) = 0$$

where $f, g \in \mathbb{Z}[x]$ are the monic polynomials defined by $f(x) = x$ and $g(x) = x - 1$. Let $\alpha, \beta \in \mathcal{O}_L$. Let f_α and f_β be the minimal polynomials of α and β, respectively. Let $d = \deg f_\alpha$ and $e = \deg f_\beta$. We can write

$$f_\alpha(x) = x^d + c_{d-1}x^{d-1} + \cdots + c_1x + c_0,$$

for some $c_0, \ldots, c_d \in \mathbb{Z}$. Let $g \in \mathbb{Z}[x]$ be defined by

$$g(x) := (-1)^d f_\alpha(-x) = (-1)^d((-x)^d + c_{d-1}(-x)^{d-1} + \cdots + c_1(-x) + c_0).$$

Then g is monic and, furthermore, $g(-\alpha) = (-1)^d f_\alpha(\alpha) = 0$. Hence, $-\alpha \in \mathcal{O}_L$. It remains to show that $\alpha\beta, \alpha + \beta \in \mathcal{O}_L$. Note firstly that $\mathbb{Z}[\alpha]$ is finitely generated. Indeed, since f_α is monic, we have that

$$\alpha^d = \sum_{i=0}^{d-1} -c_i\alpha^i \implies \alpha^d \in \mathbb{Z} + \mathbb{Z}\alpha + \cdots + \mathbb{Z}\alpha^{d-1} = \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i.$$

Now, suppose that $\alpha^k \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$ for some $k \geq d$. Then, we can write $k = d + n$ for some $n \in \mathbb{N}$. Hence,

$$\alpha^{k+1} = \alpha^{n+1}\alpha^d = \alpha^{n+1}\left(\sum_{i=0}^{d-1} -c_i\alpha^i\right) = \left(\sum_{i=0}^{d-1} -c_i\alpha^{i+(n+1)}\right) \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i,$$

as $i + (n+1) \leq (d-1) + (n+1) = k$ for all $0 \leq i \leq d-1$, and $\alpha^k \in \sum_{i=0}^{d-1} \mathbb{Z}\alpha^i$ by assumption. Hence, it follows that $\mathbb{Z}[\alpha]$ is generated by the elements $1, \alpha, \ldots, \alpha^{d-1}$ (and hence finitely generated). By a similar argument, we see that $\mathbb{Z}[\alpha, \beta]$ is finitely generated, namely by the elements $\alpha^i\beta^j$, where $0 \leq i \leq d-1$, $0 \leq j \leq e-1$. Note now that as $\mathbb{Z}[\alpha\beta] \subset \mathbb{Z}[\alpha, \beta]$, it follows that $\mathbb{Z}[\alpha\beta]$ is finitely generated. Hence, there must exist $m \in \mathbb{N}$ and some integers $c_0, \ldots, c_{m-1}$ such that

$$(\alpha\beta)^m = \sum_{i=0}^{m-1} c_i(\alpha\beta)^i.$$

In other words, $\alpha\beta$ is a zero of the monic polynomial $f \in \mathbb{Z}[x]$, defined by

$$f(x) = x^m - c_{m-1}x^m - \cdots - c_1 x - c_0,$$

and is thus an algebraic integer. A similar argument, using the fact that $\mathbb{Z}[\alpha + \beta] \subset \mathbb{Z}[\alpha, \beta]$, shows that $\alpha + \beta$ is an algebraic integer. It follows that $\mathcal{O}_L$ is a ring. □

Lemma 1.15. *Let R be an integral domain. Then:*

(1) $\deg fg = \deg f + \deg g$ *for all* $f, g \in R[x]$*;*

(2) $R[x]$ *is an integral domain.*

Proof. (1) Let $f, g \in R[x]$. Write

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad g(x) = b_m x^m + \cdots + b_1 x + b_0,$$

for some $m, n \in \mathbb{N}$ and $a_0, \ldots, a_n, b_0, \ldots, b_m \in R$, with $a_n, b_m \neq 0$. Then

$$fg(x) = a_n b_m x^{n+m} + \cdots + (a_0 b_1 + a_1 b_0)x + a_0 b_0,$$

and since R is an integral domain, it has no zero divisors, so $a_n b_m \neq 0$. Thus $\deg fg = n + m = \deg f + \deg g$.

(2) Since R is an integral domain, it is a commutative ring. Thus, by definition of the addition and multiplication operations on $R[x]$, it is also. Combining this with (1) shows that $R[x]$ has no zero-divisors. It is hence an integral domain. □

Definition 1.16. Let R be a ring. For a subset $A \subseteq R$, the *ideal generated by A* is

$$(A) = \left\{ \sum_{a \in A} r_a \cdot a : r_a \in R, \text{ only finitely many } r_a \text{ are non-zero} \right\}.$$

It is a fact that (A) does in fact define an ideal, however, we omit proof of this. If $A = \{a_1, \ldots, a_n\}$ is a finite set, we write $(A) = (a_1, \ldots, a_n)$.

Definition 1.17. Let R be a ring and let $I \triangleleft R$ be an ideal. Then I is a *principal ideal* if $I = (a)$ for some $a \in R$.

Definition 1.18. Let R be a ring. Then R is a *principal ideal domain* if it is an integral domain and every ideal $I \lhd R$ is a principal ideal.

Definition 1.19. An integral domain R is a *Euclidean domain* if there is a *Euclidean function* $\phi : R \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ such that

(1) $\phi(a \cdot b) \geq \phi(b)$ for all $a, b \neq 0$

(2) If $a, b \in R$, with $b \neq 0$, then there are $q, r \in R$ such that
$$a = b \cdot q + r,$$
and either $r = 0$ or $\phi(r) < \phi(b)$.

Lemma 1.20. *Let K be a field. Then $K[x]$ is a Euclidean domain.*

Proof. We claim that the function $\phi : K[x] \setminus \{0\} \to \mathbb{Z}_{\geq 0}$, defined by
$$\phi(f) = \deg f$$
is a Euclidean function. By Lemma 1.4, this suffices to show the desired conclusion. Condition (1) follows from Lemma 1.4. Condition (2) follows from the division algorithm for polynomials. □

Example 1.21. $\mathbb{Z}$ is a Euclidean domain. Note that $\mathbb{Z}$ is a commutative ring with no zero-divisors (and hence an integral domain). Letting $\phi : \mathbb{Z} \setminus \{0\} \to \mathbb{Z}_{\geq 0}$ be defined by $\phi(n) = |n|$, we see that the Euclidean algorithm for the integers gives the desired result.

Theorem 1.22. *Let R be a Euclidean domain. Then every ideal is principal. That is, R is a principal ideal domain.*

Proof. Let $I \lhd R$ be given. Then either $I = \{0\} = (0)$, or there exists some non-zero $a \in I$ such that
$$\phi(a) := \min\{\phi(b) : 0 \neq b \in I\}. \tag{1}$$
For any $b \in I$, we can write $b = a \cdot q + r$ for some $q, r \in R$ such that either $r = 0$ or $\phi(r) < \phi(a)$. Note then that, as I is an ideal, we have $r = b - qa \in I$. Suppose $r \neq 0$. Then we have that $\phi(r) < \phi(a)$, contradicting (1). Hence, we must have $r = 0$. Thus, $b = qa \in (a)$. As $b \in I$ was taken to be arbitrary, it follows that $I \subseteq (a)$. On the other hand, as $a \in I$, we must have $(a) \subseteq I$. It follows that $I = (a)$. □

Note that in the above proof, we could write $aq = qa$ because R must be a commutative ring.

Corollary 1.23. *Let K be a field. Then $K[x]$ is a principal ideal domain.*

Proof. Combining Lemma 1.5 and Theorem 1.2 gives the desired result. □

Corollary 1.24. *$\mathbb{Z}$ is a principal ideal domain.*

Proof. Combining Example 1.1 and Theorem 1.2 gives the desired result. □

Lemma 1.25. *Let R be a principal ideal domain. If $p \in R$ is irreducible, then it is prime.*

Proof. Let $p \in R$ be irreducible, and suppose that $p \mid a \cdot b$. Suppose further that $p \nmid a$. Consider the ideal $(p, a) \triangleleft R$. Since R is a principal ideal domain, then $(p, a) = (d)$ for some $d \in R$. Thus, $d \mid p$ and $d \mid a$. Since $d \mid p$, there exists some $q_1 \in R$ such that $p = q_1 d$. Since p is irreducible, then either d or q_1 is a unit. If q_1 is a unit, then $d = q_1^{-1} p$, and this divides a. Thus, $a = q_1^{-1} p x$ for some $x \in R$. However, this is a contradiction, since $p \nmid a$. Thus, d must be a unit. Hence, $(p, a) = (d) = R$. Hence, we have that $1_R \in (p, a)$, and thus $1_R = rp + sa$ for some $r, s \in R$. Thus,

$$b = rpb + sab.$$

Hence, as $p \mid a \cdot b$ and $p \mid p$, we have that $p \mid b$. It follows that p is prime. □

Definition 1.26. Let R be a ring. An ideal $I \triangleleft R$ is *maximal* if $I \neq R$ and for any ideal J with $I \leq J \leq R$, either $J = I$ or $J = R$.

Lemma 1.27. *Let R be a principal ideal domain. If $p \in R$ is prime, then (p) is maximal.*

Proof. Let $(p) \leq J \leq R$ for some ideal $J \triangleleft R$. Note that, as R is a principal ideal domain, we can write $J = (j)$ for some $j \in R$. Thus we have that $(p) \leq (j) \leq R$. In other words, we can write $p = rj$ for some $r \in R$. Now, since p is prime, then either $p \mid r$ or $p \mid j$. If $p \mid j$, then $j \in (p)$, and we have $(j) \leq (p) \implies (j) = (p)$, and we are done. If $p \mid r$, then $r = sp$ for some $s \in R$. Hence, as $p \in (j)$, we can write $p = rj$, and thus

$$p = spj \implies p - spj = p(1_R - sj) = 0_R \implies 1_R = sj \implies 1_R \in (j) \implies (j) = R,$$

and we are done. Note that the above step relies on the fact that R is an integral domain. □

Lemma 1.28. *Let R be a ring. Then there exists a unique homomorphism $\mathbb{Z} \to R$.*

Proof. Let $\chi : \mathbb{Z} \to R$ be defined by

$$\chi(n) := \begin{cases} 0_R & \text{if } n = 0, \\ \chi(n-1) + 1_R & \text{if } n \geq 1, \\ -\chi(-n) & \text{if } n \leq -1. \end{cases}$$

It is easy to check that χ defines a ring homomorphism $\mathbb{Z} \to R$. Now, let $\varphi : \mathbb{Z} \to R$ be any ring homomorphism. Then as φ is a ring homomorphism, we have

$$\varphi(0) = 0_R, \quad \varphi(1) = 1_R. \tag{2}$$

We proceed by induction to show that $\varphi(n) = \chi(n)$ for all $n \geq 1$. The base case holds by (2). Let $n \geq 1$ be given and suppose that $\varphi(n) = \chi(n)$. Then, as φ is a ring homomorphism, we have that

$$\varphi(n+1) = \varphi(n) + \varphi(1) = \chi(n) + 1_R = \chi(n+1).$$

Hence, by induction, $\varphi(n) = \chi(n)$ for all $n \geq 1$. Now, let $n \leq -1$. Then

$$\chi(n) = -\chi(-n) = -\varphi(-n) = -\varphi(-1)\varphi(n) = -(-1_R)\varphi(n) = \varphi(n).$$

Thus, $\varphi = \chi$, and we are done. □

Lemma 1.29. *Let K be a field. Then the only ideals of K are $\{0\}$ and K.*

Proof. Clearly, $\{0\} \lhd K$. Now, suppose that I is a non-zero ideal of K. Then it contains some non-zero element a. Hence, $a^{-1}a = 1 \in I$. Thus, $x \cdot 1 = x \in I$, for any $x \in K$. Thus, $I = K$. □

Lemma 1.30. *Let K and L be fields. Then any field homomorphism $K \to L$ is injective.*

Proof. Let $\varphi : K \to L$ be a field homomorphism. Then, as $\ker\varphi \lhd K$, we have that $\ker\varphi = \{0_K\}$, or $\ker\varphi = K$. If $\ker\varphi = K$, then $0_L = \varphi(0_K) = \varphi(1_K) = 1_L$, as φ is a ring homomorphism. However, this contradicts the fact that L is a field. Thus, we must have that $\ker\varphi = \{0_K\}$, and φ is thus injective. □

Definition 1.31. Let R be a ring. The *characteristic* $\operatorname{char}(R)$ is defined as the unique integer $n \geq 0$ such that $\ker\chi = (n)$. Recall that this exists, as $\mathbb{Z}$ is a principal ideal domain, and $\ker\chi$ is an ideal.

Example 1.32. Note that the inclusion $\mathbb{Z} \hookrightarrow \mathbb{Q}$ must be the unique ring homomorphism $\mathbb{Z} \to \mathbb{Q}$. The kernel of this inclusion is $\{0\} = (0)$. Hence, $\operatorname{char}\mathbb{Q} = 0$.

Theorem 1.33. *Let $\varphi : K \to L$ be a field homomorphism. Then* $\operatorname{char} K = \operatorname{char} L$.

Proof. Let χ_K and χ_L be the unique ring homomorphisms $\mathbb{Z} \to K$ and $\mathbb{Z} \to L$, respectively. Then

$$\begin{tikzcd} & \mathbb{Z} \arrow[dl, "\chi_K"'] \arrow[dr, "\chi_L"] & \\ K \arrow[rr, "\varphi"'] & & L \end{tikzcd}$$

commutes, as the composite $\varphi \circ \chi_K$ is a ring homomorphism $\mathbb{Z} \to L$. Thus, by Lemma 1.8, $\varphi \circ \chi_K = \chi_L$. Hence, $\ker(\varphi \circ \chi_K) = \ker(\chi_L)$. But φ is injective (by Lemma 1.10), so $\ker(\varphi \circ \chi_K) = \ker(\chi_K) = \ker(\chi_L)$. In other words, $\operatorname{char} K = \operatorname{char} L$. □

Corollary 1.34. *Let $L/\mathbb{Q}$ be a field extension. Then* $\operatorname{char} L = 0$.

Proof. The inclusion $\mathbb{Q} \hookrightarrow L$ is a field homomorphism. Combining this with Example 1.2 and Theorem 1.3 gives the desired result. □

Definition 1.35 (Formal Differentiation). Let K be a field. *Formal Differentiation* is a linear map $D : K[x] \to K[x]$ of vector spaces over K, defined by

$$D(f(x)) = D\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=1}^{n} i a_i x^{i-1} \in K[x].$$

For $f \in K[x]$, we write $D(f) = f'$.

Lemma 1.36. *Let K be a field and $f \in K[x]$. If* $\operatorname{char} K = 0$, *then* $f' = 0 \iff \deg f = 0$.

Proof. The reverse implication is trivial. We proceed with a proof of the forward implication. Let $\deg f = n \geq 1$. Then

$$f(x) = \sum_{i=0}^{n} a_i x^i, \quad f'(x) = \sum_{i=1}^{n} i a_i x^{i-1}$$

for some $a_0, \ldots, a_n$ with $a_n \neq 0$. Suppose $f' = 0$. Then (as $\operatorname{char} K = 0$) we must have $a_i = 0$ for all $i \in \{1, \ldots, n\}$. But then $\deg f = 0$, a contradiction. Thus if $f' = 0$, then $\deg f = 0$ and we are done. □

Lemma 1.37. *Let R be a ring. Let $I \triangleleft R$ be an ideal. Then I is maximal if and only if R/I is a field.*

Proof. R/I is a field if and only if $\{0\}$ and R/I are the only ideals of R/I. By the ideal correspondence, this is the same as saying that I and R are the only ideals of R that contain I. In other words, I is maximal. □

Corollary 1.38. *Let K be a field, and $f \in K[x]$ be irreducible. Then $K[x]/(f)$ is a field.*

Proof. Recall that, as K is a field, $(f) \triangleleft K[x]$ is maximal. Thus $K[x]/(f)$ is a field. □

Corollary 1.39. *Let L/K be a field extension and $\alpha \in L$ be algebraic over K. Then $K[\alpha] = K(\alpha)$.*

Proof. Let f_α be the minimal polynomial of α. Then f_α is irreducible, and $K[\alpha]/(f_\alpha)$ is hence a field. Now, consider the evaluation map $\varphi : K[x] \to K[\alpha]$, defined by $\varphi(f) = f(\alpha)$. Then, by the first isomorphism theorem for rings, we have

$$K[x]/\ker\varphi \cong \operatorname{im}\varphi = K[\alpha].$$

It follows, by Lemma 1.3, that $\ker\varphi = (f_\alpha)$. Hence, $K[x]/(f_\alpha) \cong K[\alpha]$. Thus, $K[\alpha]$ is a field, and hence $K(\alpha) \subseteq K[\alpha]$, by definition of $K(\alpha)$. But we must also have $K[\alpha] \subseteq K(\alpha)$, by definition of $K[\alpha]$. Hence, $K[\alpha] = K(\alpha)$, and we are done. □

Theorem 1.40. *Let R be a Euclidean domain, and let B be an $m \times n$ matrix with entries in R. Then B can be reduced, via elementary row and column operations, to an $m \times n$ matrix D with entries in R satisfying:*

(1) $D_{ij} = 0$ *whenever* $i \neq j$.

(2) $D_{11} \mid D_{22} \mid \cdots$.

Proof. Omitted. The proof is not too advanced, but it is fairly lengthy. □

2. Lattices

Definition 2.1. Let $n \in \mathbb{N}$. A lattice $\Lambda \subset \mathbb{R}^n$ is a subgroup of the form

$$\bigoplus_{i=1}^{n} \mathbb{Z}v_i,$$

where $\{v_1, \dots, v_n\}$ is a basis for $\mathbb{R}^n$.

Recall that subgroups of $\mathbb{R}^n$ make sense, as $\mathbb{R}^n$ is a vector space, and hence an abelian group.

Here, vol denotes the Lebesgue measure on $\mathbb{R}^n$. Recall that, in addition to the standard properties of a measure, vol satisfies:

(1) $\operatorname{vol}(E + x) = \operatorname{vol}(E)$, for any $E \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$.

(2) $\operatorname{vol}(T(E)) = |\det(T)| \cdot \operatorname{vol}(E)$, for any $E \subseteq \mathbb{R}^n$ and linear mapping $T : \mathbb{R}^n \to \mathbb{R}^n$.

(3) $\operatorname{vol}(\lambda E) = |\lambda|^n \operatorname{vol}(E)$, for any $E \subseteq \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

Note that (3) follows from (2).

Definition 2.2. Let $\Lambda \subset \mathbb{R}^n$ be a lattice. The *covolume* $A(\Lambda)$ of Λ is defined by

$$A(\Lambda) = \operatorname{vol}\left(\left\{\sum_{i=1}^{n} t_i v_i : t_i \in [0,1)\right\}\right),$$

where $\Lambda = \bigoplus_{i=1}^{n} \mathbb{Z}v_i$. The set

$$P = \left\{\sum_{i=1}^{n} t_i v_i : t_i \in [0,1)\right\}$$

is called the *fundamental parallelotope* of Λ with respect to the basis $\{v_1, \dots, v_n\}$.

Definition 2.3. Let X be a topological space. A subset $S \subseteq X$ is *discrete* if for every $s \in S$, there exists some open set $U \subseteq X$ such that $U \cap S = \{s\}$.

Theorem 2.4. *Let $f : X \to Y$ be a homeomorphism. Let $S \subset X$. If $f(S)$ is discrete, then S is discrete.*

Proof. Suppose $f(S) \subset Y$ is discrete. Then for every $s \in S$, there exists some open set $U \subseteq Y$ such that $U \cap f(S) = \{f(s)\}$. Thus,

$$f^{-1}(U \cap f(S)) = f^{-1}(\{s\}) \implies f^{-1}(U) \cap f^{-1}(f(S)) = f^{-1}(\{s\}) \implies f^{-1}(U) \cap S = \{s\}$$

Hence, as f is a homeomorphism, then f^{-1} is continuous, and thus $f^{-1}(U)$ is open in X. It thus follows that S is discrete. □

Definition 2.5. The *Euclidean norm* on $\mathbb{R}^n$ is denoted $\|\cdot\|$ and is defined by

$$\|v\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

where $v = (x_1, x_2, \cdots, x_n)$. The ℓ^∞ *norm* on $\mathbb{R}^n$ is denoted $\|\cdot\|_\infty$ and is defined by

$$\|v\|_\infty = \max_{i \in \{1,\ldots,n\}} |x_i|$$

where $v = (x_1, x_2, \cdots, x_n)$. It is a fact that these do define norms on $\mathbb{R}^n$, however, we omit proof of this.

Theorem 2.6. *Let $v \in \mathbb{R}^n$. Then $\|v\|_\infty \leq \|v\|$.*

Proof. We have

$$\|v\|_\infty = \max_{i \in \{1,\ldots,n\}} |x_i| = \max_{i \in \{1,\ldots,n\}} \sqrt{x_i^2} \leq \sqrt{\sum_{i=1}^{n} x_i^2} = \|x\|,$$

as required. □

It is important to note that the following definition of continuity of a function $f : \mathbb{R}^n \to \mathbb{R}^n$ coincides with the topological definition, when we consider $\mathbb{R}^n$ with the usual topology.

Definition 2.7. A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is *continuous* if for all $\varepsilon > 0$ and $v \in \mathbb{R}^m$, there exists $\delta > 0$ such that $\|v - w\| < \delta \implies \|f(v) - f(w)\| < \varepsilon$.

Definition 2.8. A function $f : \mathbb{R}^n \to \mathbb{R}^n$ is *Lipschitz continuous* if there exists some $L \geq 0$ such that for all $v_1, v_2 \in \mathbb{R}^m$,

$$\|f(v_1) - f(v_2)\| \leq L\|v_1 - v_2\|.$$

We call L the *Lipschitz constant* of f.

Theorem 2.9. *If a function $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous, then it is continuous.*

Proof. Let $L \geq 0$ be the Lipschitz constant of f. The case of $L = 0$ is trivial. Suppose that $L > 0$. Let $v_1, v_2 \in \mathbb{R}^n$ and let $\varepsilon > 0$. Let $\delta = \varepsilon/2L$. Then $\|v_1 - v_2\| < \delta \implies \|f(v_1) - f(v_2)\| \leq L\delta = \varepsilon/2 < \varepsilon$. Continuity follows. □

Theorem 2.10. *The inverse of any linear bijection $f : \mathbb{R}^n \to \mathbb{R}^n$ is linear.*

Proof. Let $w_1, w_2 \in \mathbb{R}^n$. Then as f is a bijection, it follows that $w_1 = f(v_1)$, $w_2 = f(v_2)$ for some $v_1, v_2 \in \mathbb{R}^n$. Hence for any $\lambda_1, \lambda_2 \in \mathbb{R}$, we have $f^{-1}(\lambda_1 w_1 + \lambda_2 w_2) = f^{-1}(\lambda_1 f(v_1) + \lambda_2 f(v_2)) = f^{-1}(f(\lambda_1 v_1 + \lambda_2 v_2)) = \lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 f^{-1}(w_1) + \lambda_2 f^{-1}(w_2)$. Linearity of f^{-1} follows. □

Theorem 2.11. *Any linear mapping $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuous.*

Proof. Let $v_1 = \sum_{i=1}^n \lambda_i e_i, v_2 = \sum_{i=1}^n \mu_i e_i \in \mathbb{R}^n$. Then

$$\begin{aligned}\|f(v_1) - f(v_2)\| = \left\| f\left(\sum_{i=1}^n (\lambda_i - \mu_i)e_i\right)\right\| &= \left\|\sum_{i=1}^n (\lambda_i - \mu_i) f(e_i)\right\| \\ &\leq \sum_{i=1}^n |\lambda_i - \mu_i| \|f(e_i)\| \\ &\leq \left(\sum_{i=1}^n \|f(e_i)\|\right) \max_{i\in\{1,\ldots,n\}} |\lambda_i - \mu_i| \\ &= L\|v_1 - v_2\|_\infty \leq L\|v_1 - v_2\|,\end{aligned}$$

where $L = \sum_{i=1}^n \|f(e_i)\| \geq 0$. Thus f is Lipschitz continuous, and hence continuous. □

Theorem 2.12. *Consider $\mathbb{R}^n$ with the usual topology. Let $X \subset \mathbb{R}^n$ be discrete and closed. Then if $K \subset \mathbb{R}^n$ is compact, the intersection $X \cap K$ is finite.*

Proof. Suppose $X \cap K$ were infinite. As K is compact, then it is closed and bounded (by the Heine-Borel Theorem). Thus $X \cap K$ is closed, as it is an intersection of two closed subsets of $\mathbb{R}^n$. Moreover, $X \cap K \subseteq K$, and thus $X \cap K$ is bounded. Hence $X \cap K$ is closed and bounded, and is thus compact, by the Heine-Borel Theorem. Hence the subspace topology on $X \cap K$ is compact. Note that as X is discrete, then $\{x\}$ is open in the subspace topology on $X \cap K$, for all $x \in X$. Moreover,

$$X \cap K \subseteq \bigcup_{x \in X\cap K} \{x\}. \tag{3}$$

Suppose $y \in \{x\}$ for some $x \in X \cap K$. Then $y = x \implies y \in X \cap K$. Hence

$$\bigcup_{x\in X\cap K} \{x\} \subseteq X \cap K \implies X \cap K = \bigcup_{x\in X\cap K} \{x\} \text{ (by (3))}.$$

Thus $\{x\}_{x\in X\cap K}$ is an open cover of the subspace topology on $X \cap K$. Hence, as the subspace topology on $X \cap K$ is compact, $\{x\}_{x\in X\cap K}$ must have a finite subcover. That is,

$$X \cap K = \bigcup_{i=1}^n \{x_i\}$$

for some $x_1, \ldots, x_n \in X \cap K$. But then $X \cap K$ must be finite, which is a contradiction. Hence $X \cap K$ cannot be infinite. □

Theorem 2.13. *Let $m \in \mathbb{N}$ and consider $\mathbb{R}^n$ with the usual topology. Then $\mathbb{Z}^n \subset \mathbb{R}^n$ is both discrete and closed.*

Proof. Let $v_1 = (k_1, k_2, \ldots, k_n), v_2 = (j_1, j_2, \ldots, j_n) \in \mathbb{Z}^n$ be given. Then

$$\|v_1 - v_2\| \geq \|v_1 - v_2\|_\infty = \max_{i\in\{1,\ldots,n\}} |k_i - j_i| \in \mathbb{Z}.$$

if $\|v_1 - v_2\|_\infty = 0$, we have $v_1 = v_2$. Thus if $v_1 \neq v_2$ for some $v_1, v_2 \in \mathbb{Z}^n$, the above implies that $\|v_1 - v_2\| \geq 1$. Now, let $v \in \mathbb{Z}^n$. Let $B_{1/2}(v)$ denote the open ball of radius 1/2, centred at v. Then, by our previous working, $B_{1/2}(v) \cap \mathbb{Z}^n = v$. It follows that $\mathbb{Z}^n$ is discrete. Now, suppose $w = (y_1, \ldots, y_n) \in \mathbb{R}^n \setminus \mathbb{Z}^n$. Then for any $v = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, we have

$$\|w - v\| \geq \|w - v\|_\infty = \max_{i \in \{1,\ldots,n\}} |y_i - x_i| \geq \max_{i \in \{1,\ldots,n\}} \max\{|y_i - \lfloor y_i \rfloor|, |\lceil y_i \rceil - y_i|\} = \delta > 0,$$

as there must exist some $i \in \{1, \ldots, n\}$ such that $y_i \notin \mathbb{Z}$. Thus, let $0 < \varepsilon < \delta$. Then $B_\varepsilon(w) \subseteq \mathbb{R}^n \setminus \mathbb{Z}^n$. It follows that $\mathbb{R}^n \setminus \mathbb{Z}^n$ is open, and hence $\mathbb{Z}^n$ is closed. □

Theorem 2.14. *Let X and Y be topological spaces, and let $f : X \to Y$ be a homeomorphism. Let $S \subseteq X$. If $f(S)$ is closed in Y, then S is closed in X.*

Proof. Note that as f is a homeomorphism, then f^{-1} is continuous. Hence if $f(S)$ is closed in Y, then $Y \setminus f(S)$ is open in Y, and hence $f^{-1}(Y \setminus f(S)) = f^{-1}(Y) \setminus f^{-1}(f(S)) = X \setminus S$ is open in X. Thus, S is closed in X. □

Theorem 2.15. *Lattices in $\mathbb{R}^n$ are discrete and closed.*

Proof. Consider $\mathbb{R}^n$ with the usual topology. This topology is induced by the Euclidean norm, and hence our previous theorems regarding linear maps and continuity are still valid. Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice. Then

$$\Lambda = \bigoplus_{i=1}^n \mathbb{Z} v_i,$$

where $v_1, \ldots, v_n$ form a basis of $\mathbb{R}^n$. Now, define a map $g : \mathbb{R}^n \to \mathbb{R}^n$ by

$$g(v) = g\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i e_i$$

where e_i is the ith standard basis vector. Let $w_1, w_2 \in \mathbb{R}^n$ and $r_1, r_2 \in \mathbb{R}$. Then writing

$$w_1 = \sum_{i=1}^n \lambda_i v_i, \quad w_2 = \sum_{i=1}^n \mu_i v_i$$

for some scalars $\lambda_1, \mu_1, \ldots, \lambda_n, \mu_n \in \mathbb{R}$, we have that

$$\begin{aligned} g(r_1 w_1 + r_2 w_2) = g\left(\sum_{i=1}^n (r_1 \lambda_i + r_2 \mu_i) v_i\right) = \sum_{i=1}^n (r_1 \lambda_i + r_2 \mu_i) e_i &= r_1 \sum_{i=1}^n \lambda_i e_i + r_2 \sum_{i=1}^n \mu_i e_i \\ &= r_1 g(w_1) + r_2 g(w_2). \end{aligned}$$

Hence g is linear. We also have that

$$g(w_1) = g(w_2) \implies \sum_{i=1}^n \lambda_i e_i = \sum_{i=1}^n \mu_i e_i \implies \lambda_i = \mu_i \text{ for all } i \in \{1, \ldots, n\} \implies w_1 = w_2.$$

So g is injective. Furthermore, given $w = \sum_{i=1}^n \lambda_i e_i \in \mathbb{R}^n$, we have $w = g\left(\sum_{i=1}^n \lambda_i v_i\right)$, and hence g is surjective. Hence g is a linear bijection. Hence g is a continuous bijection, and furthermore, it has a continuous inverse. Thus g is a homeomorphism. Moreover, we have

$$g(\Lambda) = g\left(\bigoplus_{i=1}^n \mathbb{Z} v_i\right) = \bigoplus_{i=1}^n \mathbb{Z} e_i = \mathbb{Z}^n.$$

Hence as g is a homeomorphism, then as $\mathbb{Z}^n$ is discrete and closed, it follows that Λ is discrete and closed. □

Corollary 2.16. *Let $\Lambda \subseteq \mathbb{R}^n$ be a lattice. If $K \subset \mathbb{R}^n$ is compact, then the intersection $\Lambda \cap K$ is finite.*

Theorem 2.17 (Minkowski's Theorem).

Let Λ be a lattice in $\mathbb{R}^n$, and let $E \subset \mathbb{R}^n$ be a subset satisfying the following conditions:

(1) The boundary ∂E has volume 0.

(2) E is convex.

(3) E is centrally symmetric ($x \in E \iff -x \in E$).

Then if $\operatorname{vol}(E) > 2^n A(\Lambda)$, E contains a non-zero point of Λ. If E is compact, the conclusion holds under the weaker assumption that $\operatorname{vol}(E) \geq 2^n A(\Lambda)$.

Proof. We first address the case of strict inequality. Let $\{v_1, \ldots, v_n\}$ be a $\mathbb{Z}$-basis for Λ, and let P be the fundamental parallelotope of Λ with respect to this basis. Now, since $\{v_1, \ldots, v_n\}$ is a set of n linearly independent vectors of $\mathbb{R}^n$, it is automatically a basis of $\mathbb{R}^n$. Thus, for any vector $x \in \mathbb{R}^n$, we can write

$$x = \sum_{i=1}^{n} r_i v_i$$

for some real numbers $r_1, \ldots, r_n$. We can then write $r_i = k_i + a_i$ for some $k_i \in \mathbb{Z}$ and $a_i \in [0, 1)$, for each $i \in \{1, \ldots, n\}$. Hence, we have

$$x = \sum_{i=1}^{n} r_i v_i = \sum_{i=1}^{n} (k_i + a_i) v_i = \sum_{i=1}^{n} k_i v_i + \sum_{i=1}^{n} a_i v_i = \lambda + p,$$

for some $\lambda \in \Lambda$ and $p \in P$. Now, suppose

$$(\lambda + P) \cap (\mu + P) \neq \emptyset$$

for some $\lambda, \mu \in \Lambda$ with $\lambda \neq \mu$. Then we have that $\lambda + p_1 = \mu + p_2$ for some $p_1, p_2 \in P$. Thus, $\lambda - \mu = p_2 - p_1$. Thus, we can write

$$\lambda - \mu = \sum_{i=1}^{n} a_i v_i - \sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (a_i - b_i) v_i,$$

for some $a_1, \ldots, a_n, b_1, \ldots, b_n \in [0, 1)$. Note then that $(a_i - b_i) \in (-1, 1)$ for all $i \in \{1, \ldots, n\}$. But, as $\lambda \neq \mu$, then $\lambda - \mu$ is a non-zero element of Λ. As Λ is a lattice, then this means that at least one of the $(a_i - b_i)$ terms must be a non-zero integer, and lie outside the interval $(-1, 1)$, a contradiction. Thus,

$$(\lambda + P) \cap (\mu + P) = \emptyset$$

for all $\lambda, \mu \in \Lambda$ such that $\lambda \neq \mu$. Hence, as $x \in \mathbb{R}^n$ was taken to be arbitrary, it follows that

$$\mathbb{R}^n = \bigsqcup_{\lambda \in \Lambda} (\lambda + P).$$

Thus, we can write

$$\frac{1}{2}E = \frac{1}{2}E \cap \bigsqcup_{\lambda \in \Lambda} (\lambda + P) = \bigsqcup_{\lambda \in \Lambda} \left(\frac{1}{2}E \cap (\lambda + P) \right).$$

Hence,

$$\begin{aligned} A(\Lambda) = \operatorname{vol}(P) < \frac{1}{2^n}\operatorname{vol}(E) \le \operatorname{vol}\left(\frac{1}{2}E\right) &= \operatorname{vol}\left(\bigsqcup_{\lambda\in\Lambda}\left(\frac{1}{2}E\cap(\lambda+P)\right)\right) \\ &= \sum_{\lambda\in\Lambda}\operatorname{vol}\left(\frac{1}{2}E\cap(\lambda+P)\right) \\ &= \sum_{\lambda\in\Lambda}\operatorname{vol}\left(\left(\frac{1}{2}E-\lambda\right)\cap P\right). \end{aligned}$$

Now, suppose that

$$\left(\frac{1}{2}E-\lambda\right)\cap\left(\frac{1}{2}E-\mu\right)=\emptyset$$

for all $\lambda,\mu\in\Lambda$ with $\lambda\neq\mu$. Then,

$$\operatorname{vol}(P)\ge\operatorname{vol}\left(\bigsqcup_{\lambda\in\Lambda}\left(\frac{1}{2}E-\lambda\right)\cap P\right)=\sum_{\lambda\in\Lambda}\operatorname{vol}\left(\left(\frac{1}{2}E-\lambda\right)\cap P\right),$$

a contradiction. Thus, there must exist some $\lambda,\mu\in\Lambda$ such that

$$\left(\frac{1}{2}E-\lambda\right)\cap\left(\frac{1}{2}E-\mu\right)\neq\emptyset.$$

As E is centrally symmetric and convex, this implies that $\lambda-\mu$ is a non-zero element of $\Lambda\cap E$.

Now, we must consider the case of non-strict inequality. By the Heine-Borel Theorem, E is closed and bounded. Furthermore, for any $m\in\mathbb{N}$,

$$\operatorname{vol}\left(\left(1+\frac{1}{m}\right)E\right)=\left(1+\frac{1}{m}\right)^n\operatorname{vol}(E)>\operatorname{vol}(E)=2^nA(\Lambda),$$

so we can use the first part of the proof to deduce that there exists some non-zero element $\lambda_m\in\left(1+\frac{1}{m}\right)E$, for each $m\in\mathbb{N}$. Note that each of these points is contained in $2E\cap\Lambda$ (as $1+1/m\le 2$ for all $m\in\mathbb{N}$), which is a finite set, as Λ is a lattice and E is compact. Hence, by the pigeonhole principle, there must exist some non-zero $\lambda\in\Lambda$ such that

$$\lambda\in\bigcap_{m\in\mathbb{N}}\left(1+\frac{1}{m}\right)E=E,$$

and we are done. □

Note that condition (1) is necessary to invoke additivity of the Lebesgue measure. If the volume of the boundary of E were non-zero, we would not necessarily be able to partition E into a disjoint union of tilings as we did in the proof.

Definition 2.18. Let G be a group. A *torsion element* is an element $g\in G$ of finite order. A group G is called *torsion-free* if the only torsion element of G is the identity element.

Lemma 2.19. *$\mathbb{Z}^n$ is finitely generated and torsion-free, for all $n\ge 0$.*

Proof. The $n = 0$ case is trivial (we just get the trivial group). Let $n \geq 1$. Then $\mathbb{Z}^n = \langle e_1, \ldots, e_n \rangle$, where e_i is the element whose entries are all zero except for the ith entry, which is one. Let $(a_1, \ldots, a_n) \in \mathbb{Z}^n$. Then, for any $k \geq 1$, $k(a_1, \ldots, a_n) = 0 \iff ka_i = 0$ for all $i \in \{1, \ldots n\} \iff a_i = 0$ for all $i \in \{1, \ldots n\}$. Hence, the only element of $\mathbb{Z}^n$ of finite order is the identity element. □

Lemma 2.20. *Every subgroup of a torsion-free group is torsion-free.*

Proof. Let G be torsion-free and $H \leq G$. Suppose that H has an element of finite order. Then there exist $h \in H$ and $n \geq 1$ such that $h^n = e_H = e_G$. Since $h \in G$, this is a contradiction. □

Lemma 2.21 (Sandwich Lemma).

(1) Let $H \subset G$ be abelian groups such that $G \cong \mathbb{Z}^n$ for some $n \geq 1$. Then $H \cong \mathbb{Z}^m$ for some $m \leq n$.

(2) Let $K \subset H \subset G$ be abelian groups such that $K \cong \mathbb{Z}^n$ and $G \cong \mathbb{Z}^n$ for some $n \geq 1$. Then $H \cong \mathbb{Z}^n$.

(3) Let $H \subset G$ be abelian groups such that $H \cong G \cong \mathbb{Z}^n$. Then G/H is finite.

Proof. (1) Note firstly that H is finitely generated, abelian, and torsion-free, as it is a subgroup of $\mathbb{Z}^n$ for some $n \geq 1$. By the Fundamental Theorem of Finitely Generated Abelian Groups, we have

$$H \cong \mathbb{Z}/r_1\mathbb{Z} \oplus \mathbb{Z}/r_2\mathbb{Z} \oplus \cdots \times \mathbb{Z}/r_k\mathbb{Z} \oplus \mathbb{Z}^m,$$

for some $k, m \in \mathbb{N}$ and non-zero integers $r_1, \ldots, r_k$ such that $r_1 \mid r_2 \mid \cdots \mid r_k$. Note that if $k \neq 0$, then this contradicts the fact that H is torsion-free. Thus, $H \cong \mathbb{Z}^m$ for some $m \geq 0$. We must now show why $m \leq n$. Note firstly that H is an abelian subgroup of G, and is thus a normal subgroup of G. Hence, we can conisder the quotient group G/H. Furthermore, G/H is finitely generated and abelian. Thus, we can write

$$G/H \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \mathbb{Z}/d_2\mathbb{Z} \oplus \cdots \times \mathbb{Z}/d_s\mathbb{Z} \oplus \mathbb{Z}^\ell,$$

for some $\ell, s \in \mathbb{N}$ and non-zero integers $d_1, \ldots, d_s$ such that $d_1 \mid d_2 \mid \cdots \mid d_s$. Let p be a prime such that $p \nmid d_i$ for any $i \in \{1, \ldots, s\}$. Consider the map $\phi : G/H \to G/H$ defined by $\phi(g + H) = p(g + H)$. One can check that this map is well-defined and, moreover, is a homomorphism. Furthermore, it is injective, as $\mathbb{Z}^\ell$ is torsion free, and the order of any element of $\mathbb{Z}/d_i\mathbb{Z}$ must divide d_i, for each $i \in \{1, \ldots, s\}$, so the kernel must be trivial. Now, consider the map $\varphi : H/pH \to G/pG$, defined by $\varphi(h + pH) = h + pG$. Again, one can check that this map is well-defined and a homomorphism. Clearly, $pH \in \ker\varphi$. Note that

$$h + pH \in \ker\varphi \iff h + pG = pG \iff h = pg \text{ for some } g \in G.$$

Note that if $h = pg$ for some $g \in G$, then

$$\begin{aligned} p(g + H) = pg + H = h + H = H &\implies g + H \in \ker\phi \\ &\implies g + H = H \iff g \in H. \end{aligned}$$

Hence, $h = pg \in pH \implies h + pH = pH$. Thus, $\ker\varphi = \{pH\}$. Hence, φ is injective. Note that $[\mathbb{Z}^m : p\mathbb{Z}^m] = p^m$, as for an element $(a_1, \ldots, a_m) \in \mathbb{Z}^m$, we have p possible remainders modulo p for each entry. Similarly, $[\mathbb{Z}^n : p\mathbb{Z}^n] = p^n$. Thus, $p^m = |H/pH| \leq |G/pG| = p^n \implies m \leq n$, as was to be shown.

(2) follows directly from (1).

(3) Again, by the Fundamental Theorem of Finitely Generated Abelian Groups, we have that $G/H \cong \mathbb{Z}^a \oplus T$, where T is a finite abelian group. Again, let p be a prime that does not divide $|T|$. By an argument similar to that of the proof of (1), we see that the map $\phi : G/H \to G/H$ defined by $\phi(g + H) = p(g + H)$ is an injective homomorphism. Now, define $\varphi : G/pG \to G/(H + pG)$ by $\varphi(g + pG) = g + (H + pG)$. One can easily check that this defines a surjective homomorphism. Furthermore, $\ker \varphi = \{g + pG : g \in H + pG\} = (H + pG)/pG$. Now, we also have a map $\pi : H \to (H + pG)/pG$, defined by $\pi(h) = h + pG$. Again, one can easily check that this defines a surjective homomorphism. Furthermore, $\ker \pi = pH$. Thus, by the first isomorphism theorem, we have $(H + pG)/pG \cong H/pH$ and, moreover,

$$(G/pG)/((H + pG)/pG) \cong G/(H + pG).$$

Since $|H/pH| = |G/pG| = p^n$, we have

$$|G/(H + pG)| = |(G/pG)/((H + pG)/pG)| = |(G/pG)|/|(H/pH)| = 1.$$

But, if $a > 0$, we have by the third isomorphism theorem that

$$|G/(H + pG)| = |(G/H)/p(G/H)| = |(\mathbb{Z}/p\mathbb{Z})^a \oplus T/pT| = |(\mathbb{Z}/p\mathbb{Z})^a| = p^a > 1,$$

a contradiction. Thus, we must have $a = 0$, and the result follows. (Note that T/pT vanishes, as $p \nmid |T|$, so $pT = T$ here). □

3. Number Fields

Definition 3.1. A *number field* is a finite field extension over $\mathbb{Q}$.

Definition 3.2. Let L be a number field. A *complex embedding* of L is a field homomorphism $\sigma : L \to \mathbb{C}$.

Lemma 3.3. *Let L be a number field, and let $\alpha \in \mathcal{O}_L$. Then f_α is irreducible.*

Proof. Suppose we can write $f_\alpha(x) = g(x)h(x)$ for some non-constant $g, h \in L[x]$. Note that Lemma 1.4 implies that either $\deg g, \deg h < \deg f_\alpha$. Note then that

$$f_\alpha(\alpha) = 0 \implies g(\alpha)h(\alpha) = 0.$$

As L is a field, then it has no zero-divisors. Hence $g(\alpha) = 0$ or $h(\alpha) = 0$. But this contradicts the fact that f_α is the minimal polynomial of α. It follows that f_α is irreducible. □

Lemma 3.4. *Let $L/\mathbb{Q}$ be a field extension. Then* $\operatorname{char} L = 0$.

Proof. The inclusion $\mathbb{Q} \hookrightarrow L$ is a field homomorphism. Combining this with Example 1.2 and Theorem 1.3 gives the desired result. □

Theorem 3.5. *Let L be a number field and $\alpha \in \mathcal{O}_L$. Then f_α and f'_α generate the unit ideal in $L[x]$.*

Proof. Recall that f_α is irreducible. Hence, as $L[x]$ is a principal ideal domain, then f_α is prime. Thus (f_α) is maximal. Recall also that $\deg f'_\alpha = \deg f_\alpha - 1$. Thus we have that $f_\alpha \nmid f'_\alpha$ (this follows by Lemma 1.4). Now, suppose that $(f_\alpha, f'_\alpha) \neq L[x]$. Then as $(f_\alpha) \leq (f_\alpha, f'_\alpha) \leq L[x]$, it follows that $(f_\alpha, f'_\alpha) = (f_\alpha) \implies f'_\alpha \in (f_\alpha)$. However, this is a contradiction, as $f_\alpha \nmid f'_\alpha$. Hence, we must have $(f_\alpha, f'_\alpha) = L[x] = (1)$. □

Lemma 3.6. *Let L/K be an extension of number fields of degree $[L : K] = n$, and $\alpha \in L \setminus K$. Then α is algebraic over K, and* $\deg f_\alpha = [K(\alpha) : K]$.

Proof. Note firstly that, as $[L : K] = n$, then the elements $1, \alpha, \ldots, \alpha^n$ of L must be linearly dependent. That is, there must exist some $k_0, \ldots, k_n \in K$ (not all zero) such that $k_0 + \sum_{i=1}^{n} k_i\alpha^i = 0$. Thus, let $f \in K[x]$ be defined by $f(x) = k_n x^n + \cdots + k_0$. Now, let $i = \max\{0 \leq j \leq n : k_j \neq 0\}$. Then, $k_i^{-1} f \in K[x]$ defines a monic polynomial that vanishes at α. Hence, α is algebraic over K. Now, let f_α be the minimal polynomial of α. Write $d = \deg f_\alpha$. Suppose that there exist some $k_0, \ldots, k_{d-1} \in K$ (not all zero) such that $k_0 + \sum_{i=1}^{d-1} k_i\alpha^i = 0$. Then, by the same process that we used before, we can obtain a monic polynomial $p \in K[x]$ of degree $d-1$ satisfying $p(\alpha) = 0$. However, this contradicts minimality of f_α. Hence, $1, \alpha, \ldots, \alpha^{d-1}$ are linearly independent elements of $K(\alpha)$. Recall that any element of $K(\alpha)$ is of the form $p(\alpha)q(\alpha)^{-1}$ for some $p, q \in K[x]$ such that $q(\alpha) \neq 0$. Thus, by the Euclidean algorithm for polynomials, we can write $p = f_\alpha r + s$ for some $r, s \in K[x]$ with $\deg s < \deg f_\alpha$. Note then that $p(\alpha) = s(\alpha)$. Applying the same argument to q, we see that $q(\alpha) = s'(\alpha)$ for some $s' \in K[x]$ such that $\deg s' < \deg f_\alpha$ and $s'(\alpha) \neq 0$. It follows that $p(\alpha)q(\alpha)^{-1}$ is a K-linear combination of the terms $1, \alpha, \ldots, \alpha^{d-1}$. Hence, $1, \alpha, \ldots, \alpha^{d-1}$ span $K(\alpha)$, and are hence a basis of $K(\alpha)$ over K. Thus, $d = [K(\alpha) : K] = \deg f_\alpha$. □

Theorem 3.7. *Let L/K be an extension of number fields, and let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Then the number of distinct embeddings $\sigma : L \to \mathbb{C}$ such that $\sigma|_K = \sigma_0$ is equal to the degree $[L : K]$.*

Proof. We proceed by induction on $[L : K]$. If $[L : K] = 1$, then by Lemma 1.1, $L = K$ and we are done. Now, assume that $[L : K] = n > 1$, and that the statement holds for all $k < n$. As $[L : K] \neq 1$, it follows by Lemma 1.1 that there exists some $\alpha \in L \setminus K$. Thus, we have a tower

$$K \hookrightarrow K(\alpha) \hookrightarrow L$$

of field extensions. Hence, by the tower law, we have

$$[L : K] = [L : K(\alpha)][K(\alpha) : K].$$

Furthermore, $\alpha \notin K$, so $K(\alpha) \neq K$. Thus, $[K(\alpha) : K] \geq 2$, by Lemma 1.1. Hence,

$$[L : K(\alpha)] = \frac{[L : K]}{[K(\alpha) : K]} \leq \frac{[L : K]}{2} < [L : K].$$

Suppose $[K(\alpha) : K] < [L : K]$. Let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Then, by the assumption in the inductive step, we have that there are exactly $[K(\alpha) : K]$ embeddings $\tau : K(\alpha) \to \mathbb{C}$ such that $\tau|_K = \sigma_0$. Furthermore, given any such τ, there are exactly $[L : K(\alpha)]$ embeddings $\mu : L \to \mathbb{C}$ such that $\mu|_{K(\alpha)} = \tau$. Thus, there are $[K(\alpha) : K][L : K(\alpha)] = [L : K]$ embeddings $\sigma : L \to \mathbb{C}$ such that $\sigma|_K = \sigma_0$. Thus, the statement holds when $k = n$ and, by induction, we are done.

By the above, we have reduced the proof to the case of $L = K(\alpha)$. Let $\phi : K[x] \to L$ be defined by $\phi(f) = f(\alpha)$. It is clear, by the definitions of the addition and multiplication operations on $K[x]$, that ϕ is a ring homomorphism. Thus, by the first isomorphism theorem for rings, we have the isomorphism $K[x]/\ker\phi \cong \operatorname{im}\phi$. Furthermore, by Lemma 1.3, $f \in \ker\phi \iff f \in (f_\alpha)$, and $\operatorname{im}\phi = K[\alpha] = K(\alpha) = L$ (by Corollary 1.6). Thus, the map $K[x]/(f_\alpha) \to L$, $x \mapsto \alpha$ is an isomorphism. Now, since $L = K(\alpha) = K[\alpha]$, then any embedding $\sigma : L \to \mathbb{C}$ extending σ_0 is wholly determined by $\sigma(\alpha)$, as each element of L is of the form $f(\alpha)$ for some $f \in K[x]$. Furthermore,

$$0 = \sigma(f_\alpha(\alpha)) = \sigma\left(\sum_{i=0}^{\deg f_\alpha} b_i\alpha^i\right) = \sum_{i=0}^{\deg f_\alpha} \sigma(b_i)(\sigma(\alpha))^i = \sum_{i=0}^{\deg f_\alpha} \sigma_0(b_i)(\sigma(\alpha))^i = (\sigma_0 f_\alpha)(\sigma(\alpha)),$$

where $\sigma_o f_\alpha \in \mathbb{C}[x]$ is defined by

$$\sigma_0 f_\alpha(x) = \sum_{i=0}^{\deg f_\alpha} \sigma_0(b_i)(x)^i.$$

The above follows as σ is a field homomorphism, and $b_i \in K$ for all $i \in \{0, \ldots, \deg f_\alpha\}$. Now, the above shows also that $\sigma(\alpha)$ is a root of $\sigma_0 f_\alpha$. Conversely, if we take any root $\beta \in \mathbb{C}$ of $\sigma_0 f_\alpha$ then the assignment $\alpha \mapsto \beta$ extends uniquely to a field homomorphism $\sigma : L \to \mathbb{C}$, as for any $\gamma \in L$, we have

$$\sigma(\gamma) = \sigma\left(\sum_{i=0}^{n} c_i\alpha^i\right) = \sum_{i=0}^{n} \sigma_0(c_i)\beta^i,$$

for some $n \in \mathbb{N}$. Thus, the possible extensions of σ_0 are in a one-to-one correspondence with the roots of $\sigma_0 f_\alpha$. Now, recall that f_α and f'_α generate the unit ideal in $K[x]$. Thus, there exist $u, v \in K[x]$ such that $uf_\alpha + vf'_\alpha = 1$. Suppose now that $f_\alpha(\delta) = f'_\alpha(\delta)$ for some $\delta \in L$. Then

$$(uf_\alpha)(\delta) + (vf'_\alpha)(\delta) = u(\delta)f_\alpha(\delta) + v(\delta)f'_\alpha(\delta) = 0 \neq 1,$$

as K is a field. Thus, f_α and f'_α have no roots in common. Note that this in turn implies that $\sigma_0 f_\alpha$ and $(\sigma_0 f_\alpha)'$ have no roots in common. Thus, $\sigma_0 f_\alpha$ has $\deg f_\alpha$ distinct roots. Hence, there are $\deg f_\alpha = [K(\alpha) : K] = [L : K]$ such embeddings. □

Corollary 3.8. *Let L be a number field of degree n. Then there are $[L : \mathbb{Q}] = n$ complex embeddings $L \to \mathbb{C}$.*

Proof. This follows from the above lemma, with the inclusion $\sigma_0 : \mathbb{Q} \hookrightarrow \mathbb{C}$. □

If L is a number field and $\sigma : L \to \mathbb{C}$ is a complex embedding, we define $\overline{\sigma} : L \to \mathbb{C}$ by $\overline{\sigma}(\alpha) = \overline{\sigma(\alpha)}$. One can easily check that this then defines a complex embedding $\overline{\sigma} : L \to \mathbb{C}$. There are then two possibilities: $\overline{\sigma} = \sigma$, whereby σ takes values in $\mathbb{R}$, or $\overline{\sigma} \neq \sigma$. We write r for the number of embeddings $\sigma : L \to \mathbb{R}$, and s for the number of pairs $\sigma, \overline{\sigma} : L \to \mathbb{C}$ of embeddings with $\sigma \neq \overline{\sigma}$. By the above corollary, we have $r + 2s = [L : \mathbb{Q}]$.

Throughout the rest of this section, L is a number field of degree n, and $\sigma_1, \dots, \sigma_n$ denote the n complex embeddings of L.

Definition 3.9. Let $\alpha_1, \dots, \alpha_n$ be elements of L. The *discriminant* $\operatorname{disc}(\alpha_1, \dots, \alpha_n)$ of the elements $\alpha_1, \dots, \alpha_n$ is defined as

$$\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det(D)^2,$$

where D is the $n \times n$ matrix defined by $D_{ij} = \sigma_i(\alpha_j)$.

It is worth noting that, as the square is included, the value of the discriminant is invariant under reordering of the elements. Indeed, certain re-orderings can induce a sign-flip of the determinant, but the square clears this difference.

Definition 3.10. Let L/K be an extension of number fields, and let $\alpha \in L$. Consider the linear map $m_\alpha : L \to L$, given by

$$m_\alpha(\ell) = \alpha\ell,$$

for all $\ell \in L$. The *norm* of α is defined as:

$$N_{L/K}(\alpha) = \det m_\alpha,$$

and the *trace* of α is defined as follows:

$$\operatorname{tr}_{L/K}(\alpha) = \operatorname{tr} m_\alpha.$$

Note that linearity of $\operatorname{tr}_{L/K}$ and the fact that $N_{L/K}$ is multiplicative follow as a result of the properties of the determinant and the trace.

Lemma 3.11. *Let L/K be an extension of number fields, and let $\alpha \in L$. Then we have that $\operatorname{tr}_{L/K}(\alpha) = [L : K(\alpha)] \operatorname{tr}_{K(\alpha)/K}(\alpha)$ and that $N_{L/K}(\alpha) = N_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]}$.*

Proof. Recall, from the proof of the tower law, that $[L : K] = [L : K(\alpha)][K(\alpha) : K]$. Write $d = [L : K]$, $\ell = [L : K(\alpha)]$ and $p = [K(\alpha) : K]$, and let $\{v_1, \dots, v_\ell\}$ and $\{e_1, \dots, e_p\}$ define bases of L over $K(\alpha)$ and $K(\alpha)$ over K, respectively. Recall (again, from the proof of the tower law) that $\{e_j v_i : 1 \leq i \leq \ell, 1 \leq j \leq p\}$ is a basis of L over K. We can write $\alpha e_i = \sum_{j=1}^{p} b_{ji} e_j$ for some $b_{1i}, \dots, b_{pi} \in K$, for each $i \in \{1, \dots, p\}$. Let $A \in M_p(K)$ be the matrix defined by $A_{ij} = b_{ji}$. Then A is the matrix corresponding to the linear map $m_{\alpha|K(\alpha)}$, as for any $w = (w_1, \dots, w_p) \in K^p$,

$$Aw = \left(\sum_{j=1}^{p} b_{j1} w_j, \dots, \sum_{j=1}^{p} b_{jp} w_j \right)^T = (\alpha e_1) \cdot w + \cdots + (\alpha e_p) \cdot w = \alpha w.$$

Hence, $\det A = N_{K(\alpha)/K}(\alpha)$, $\operatorname{tr} A = \operatorname{tr}_{K(\alpha)/K}(\alpha)$. Let $\tilde{A} \in M_d(K)$ be the block matrix defined by

$$\tilde{A} = \underbrace{\begin{pmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & A \end{pmatrix}}_{\ell \text{ times}}.$$

Note, for any $w = (w_1, \dots, w_\ell) \in L$, we can write

$$w = \sum_{i=1}^{\ell} \sum_{j=1}^{p} w_{ij} e_j v_i = \sum_{i=1}^{\ell} \left(\sum_{j=1}^{p} w_{ij} e_j \right) v_i = (w^{(1)}, \dots, w^{(\ell)}),$$

where $w^{(i)} = (w_{i1}, \dots, w_{ip}) \in K^p$. Thus we have

$$\begin{aligned} \tilde{A} w = \sum_{i=1}^{\ell} \left(\sum_{j=1}^{p} \left(\sum_{k=1}^{p} b_{jk} w_k^{(i)} \right) e_j \right) v_i &= \sum_{i=1}^{\ell} \left(\sum_{k=1}^{p} \left(\sum_{j=1}^{p} b_{jk} e_j \right) w_k^{(i)} \right) v_i \\ &= \sum_{i=1}^{\ell} \left(\sum_{k=1}^{p} (\alpha e_k) \, w_k^{(i)} \right) v_i \\ &= \sum_{i=1}^{\ell} \alpha w^{(i)} v_i = \alpha \sum_{i=1}^{\ell} w^{(i)} v_i = \alpha w. \end{aligned}$$

Thus, $\tilde{A}$ is the matrix corresponding to the linear map m_α. Hence,

$$\begin{aligned} \det \tilde{A} &= \det(A)^\ell = N_{K(\alpha)/K}(\alpha)^{[L:K(\alpha)]} = N_{L/K}(\alpha), \\ \operatorname{tr} \tilde{A} &= \ell \operatorname{tr}(A) = [L : K(\alpha)] N_{K(\alpha)/K}(\alpha) = \operatorname{tr}_{L/K}(\alpha), \end{aligned}$$

as was to be shown. □

Lemma 3.12. *Let L/K be an extension of number fields of degree $[L : K] = \ell$, and let $\sigma_0 : K \to \mathbb{C}$ be a complex embedding. Let $\sigma_1, \dots, \sigma_\ell$ be the distinct complex embeddings such that $\sigma_{i|K} = \sigma_0$ for each $i \in \{1, \dots, \ell\}$. Then, for each $\alpha \in L$, we have*

$$\sigma_0(\operatorname{tr}_{L/K}(\alpha)) = \sum_{i=1}^{\ell} \sigma_i(\alpha), \quad \sigma_0(N_{L/K}(\alpha)) = \prod_{i=1}^{\ell} \sigma_i(\alpha).$$

Proof. We first treat the case of $L = K(\alpha)$. Let $\chi_{m_\alpha} \in K[x] \subset L[x]$ denote the characteristic polynomial of m_α. By the Cayley-Hamilton theorem, we know that $\chi_{m_\alpha}(m_\alpha) = 0$. Hence, $(\chi_{m_\alpha}(m_\alpha))(\beta) = 0$ for any $\beta \in L$. Hence, $(\chi_{m_\alpha}(m_\alpha))(1) = \chi_{m_\alpha}(\alpha) = 0$. Thus, $f_\alpha \mid \chi_{m_\alpha}$. Furthermore, $\deg \chi_{m_\alpha} = \ell = [K(\alpha) : K] = \deg f_\alpha$, and hence $\chi_{m_\alpha} = f_\alpha$. Recall that

$$\chi_{m_\alpha}(x) = x^\ell - \operatorname{tr}(m_\alpha)x^{\ell-1} + \cdots + (-1)^\ell \det(m_\alpha).$$

Hence, by the above working, $\operatorname{tr}_{L/K}(\alpha) = -a_{\ell-1}$ and $N_{L/K}(\alpha) = (-1)^\ell a_0$, where

$$f_\alpha(x) = x^\ell + a_{\ell-1}x^{\ell-1} + \cdots + a_0.$$

Furthermore, we have that $\sigma_0 f_\alpha \in \mathbb{C}[x]$, and that $\sigma_0 f_\alpha(\sigma_i(\alpha)) = \sigma_0(\sigma_i(f_\alpha(\alpha))) = 0$, for each $i \in \{1, \dots, \ell\}$. As $\mathbb{C}$ is algebraically closed, we can then write

$$\sigma_0 f_\alpha(x) = (x - \sigma_1(\alpha)) \cdots (x - \sigma_\ell(\alpha)).$$

Hence, $\sigma_0(a_0) = \sigma_0 f_\alpha(0) = (-1)^\ell \prod_{i=1}^\ell \sigma_i(\alpha)$. Furthermore, by our previous working, we also have that $\sigma_0(N_{L/K}(\alpha)) = (-1)^\ell \sigma_0(a_0) = \prod_{i=1}^\ell \sigma_i(\alpha)$. Finally, we have that

$$\sigma_0 f_\alpha(x) = \prod_{i=1}^{\ell}(x - \sigma_i(\alpha)) = x^\ell - \left(\sum_{i=1}^{\ell} \sigma_i(\alpha)\right) x^{\ell-1} + \text{l.o.t.}$$

Matching coefficients with $\sigma_o f_\alpha(x) = x^\ell + \sigma_0(a_{\ell-1})x^{\ell-1} + \cdots + \sigma_0(a_0)$ gives the desired result.

We now treat the general case. Write $[L : K] = \ell, [L : K(\alpha)] = \ell_1, [K(\alpha) : K] = \ell_2$. By the previous lemma, and our above working,

$$\begin{aligned}\sigma_0(\operatorname{tr}_{L/K}(\alpha)) = \sigma_0([L : K(\alpha)] \operatorname{tr}_{K(\alpha)/K}(\alpha)) &= [L : K(\alpha)]\sigma_0(\operatorname{tr}_{K(\alpha)/K}(\alpha)) \\ &= [L : K(\alpha)] \sum_{i=1}^{\ell_2} \sigma_i(\alpha).\end{aligned}$$

Recall that each σ_i is a complex embedding $K(\alpha) \to \mathbb{C}$ such that $\sigma_{i|K} = \sigma_0$. Recall that, for each σ_i, there are then $[L : K(\alpha)]$ embeddings $\tau : L \to \mathbb{C}$ such that $\tau_{|K(\alpha)} = \sigma_i$. This says that $\tau(\alpha) = \sigma_i(\alpha)$ for each of these embeddings, for each σ_i. Write τ_{ij} for the jth such embedding for σ_i, where $j \in \{1, \dots, \ell_1\}$. Thus $\#\{\tau_{ij} : 1 \leq i \leq \ell_1, 1 \leq j \leq \ell_2\} = \ell_1\ell_2 = \ell$, and this is thus the complete set of embeddings $L \to \mathbb{C}$ whose restriction to K gives the map σ_0. Then we have that

$$[L : K(\alpha)] \sum_{i=1}^{\ell_2} \sigma_i(\alpha) = \sum_{i=1}^{\ell_2} \ell_1 \sigma_i(\alpha) = \sum_{i=1}^{\ell_2} \left(\sum_{j=1}^{\ell_1} \tau_{ij}(\alpha) \right) = \sum_{k=1}^{\ell} \gamma_k(\alpha),$$

where $\gamma_1, \dots, \gamma_\ell$ are the distinct embeddings $L \to \mathbb{C}$ such that $\gamma_{i|K} = \sigma_0$, for each $i \in \{1, \dots, \ell\}$. A similar argument for the norm, in which the power distributes across the product, gives us the result in full. □

The case where $K = \mathbb{Q}$ tells us that

$$\operatorname{tr}_{L/K}(\alpha) = \sum_{i=1}^{n} \sigma_i(\alpha), \quad N_{L/K}(\alpha) = \prod_{i=1}^{n} \sigma_i(\alpha).$$

for any $\alpha \in L$.

Corollary 3.13. *Assume the same set-up as the previous lemma. If $\alpha \in \mathcal{O}_L$, then we have that $\operatorname{tr}_{L/K}(\alpha), N_{L/K}(\alpha) \in \mathcal{O}_K$.*

Proof. Let $\beta \in K$, and note that $f(\sigma_0(\beta)) = 0 \iff \sigma_0(\beta) = 0$. Thus, $\beta \in \mathcal{O}_K$ if and only if $\sigma_0(\beta) \in \mathcal{O}_{\mathbb{C}}$. Now, let $\alpha \in \mathcal{O}_L$. The expressions for $\sigma_0(\operatorname{tr}_{L/K}(\alpha))$ and $\sigma_0(N_{L/K}(\alpha))$ derived in the previous lemma give the desired result, as $\mathcal{O}_K$ is a ring. □

Recall that $\mathcal{O}_{\mathbb{C}}$ makes sense; we defined the ring of algebraic integers for any field extension $L/\mathbb{Q}$.

Corollary 3.14. *Let $\alpha \in \mathcal{O}_L$. Then $\operatorname{tr}_{L/\mathbb{Q}}(\alpha), N_{L/\mathbb{Q}}(\alpha) \in \mathbb{Z}$.*

Proof. This follows by the previous lemma, as $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$. □

Lemma 3.15. *Let $\alpha_1, \dots, \alpha_n$ be elements of L. Then $\operatorname{disc}(\alpha_1, \dots, \alpha_n) = \det(T)$, where T is the $n \times n$ matrix defined by $T_{ij} = \operatorname{tr}_{L/\mathbb{Q}}(\alpha_i \alpha_j)$.*

Proof. Note that

$$T_{ij} = \sum_{k=1}^{n} \sigma_k(\alpha_i)\sigma_k(\alpha_j) = \sum_{k=1}^{n} D_{ki} D_{kj} = (D^T D)_{ij}.$$

Thus, $\det(T) = \det\bigl(D^T D\bigr) = \det\bigl(D^T\bigr)\det(D) = \det(D)^2 = \operatorname{disc}(\alpha_1, \dots, \alpha_n)$. □

Lemma 3.16. *Let $\alpha \in \mathcal{O}_L^\times$. Then $N_{L/\mathbb{Q}}(\alpha) = \pm 1$.*

Proof. Firstly, suppose that $\alpha \in \mathcal{O}_L^\times$. Then, there exists $\beta \in \mathcal{O}_L$ such that $\alpha\beta = 1$. As the norm is multiplicative, we have that $1 = N_{L/\mathbb{Q}}(\alpha\beta) = N_{L/\mathbb{Q}}(\alpha)N_{L/\mathbb{Q}}(\beta)$. As $\mathbb{Z}^\times = \{\pm 1\}$, Corollary 3.3 implies that $N_{L/\mathbb{Q}}(\alpha) = \pm 1$. □

Definition 3.17. Let F be a field and V a vector space over F. A bilinear form $U : V \times V \to F$ is *non-degenerate* if

$$U(w, v) = 0 \text{ for all } w \in V \implies v = 0,$$

for any $v \in V$.

Lemma 3.18. *Let $\alpha_1, \dots, \alpha_n$ be a basis of L over $\mathbb{Q}$. Then $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$ if and only if the bilinear form $U : L \times L \to \mathbb{Q}$ defined by*

$$U(\alpha, \beta) = \operatorname{tr}_{L/\mathbb{Q}}(\alpha\beta)$$

is non-degenerate.

Proof. Assume that $\operatorname{disc}(\alpha_1, \dots, \alpha_n) \neq 0$. That is, $\det(T) \neq 0$. Equivalently, T is invertible. Let $\beta \in L$ and suppose that $U(\gamma, \beta) = 0$ for all $\gamma \in L$. We can write $\beta = \sum_{i=1}^{n} q_i \alpha_i$ for some

$q_1, \ldots, q_n \in \mathbb{Q}$. Write $Q = (q_1, \ldots, q_n)^T$. Then

$$\begin{aligned} 0 = U(\alpha_j, \beta) = U\left(\alpha_j, \sum_{i=1}^n q_i\alpha_i\right) = \sum_{i=1}^n q_i U(\alpha_j, \alpha_i) &= \sum_{i=1}^n q_i \operatorname{tr}_{L/\mathbb{Q}}(\alpha_j\alpha_i) \\ &= \sum_{i=1}^n \operatorname{tr}_{L/\mathbb{Q}}(\alpha_j\alpha_i) q_i = (T^TQ)_j, \end{aligned}$$

for any $j \in \{1, \ldots, n\}$. Thus, $T^TQ = 0$. Since T is invertible, so too is T^T, and this implies that $Q = 0$. That is, $q_i = 0$ for all $i \in \{1, \ldots, n\}$, and thus $\beta = 0$. Thus, U is non-degenerate. Now, suppose that U is non-degenerate. Suppose that T is not invertible. Then there exists some non-zero vector $Q = (q_1, \ldots, q_n) \in \mathbb{Q}^n$ such that $T^TQ = 0$. Define $\beta \in L \setminus \{0\}$ by

$$\beta = \sum_{j=1}^n q_j\alpha_j.$$

Let $\gamma \in L$ be given, and write $\gamma = \sum_{j=1}^n p_j\alpha_j$ for some $p_1, \ldots, p_n \in \mathbb{Q}$. Then

$$\begin{aligned} U(\gamma, \beta) = U\left(\sum_{j=1}^n p_j\alpha_j, \sum_{j=1}^n q_j\alpha_j\right) = \sum_{j=1}^n p_j U\left(\alpha_j, \sum_{k=1}^n q_k\alpha_k\right) &= \sum_{\substack{j=1 \\ k=1}}^n p_j q_k U(\alpha_j, \alpha_k) \\ = \sum_{\substack{j=1 \\ k=1}}^n p_j q_k \operatorname{tr}_{L/\mathbb{Q}}(\alpha_j\alpha_k) &= \sum_{j=1}^n p_j \left(\sum_{k=1}^n \operatorname{tr}_{L/\mathbb{Q}}(\alpha_j\alpha_k) q_k\right). \end{aligned}$$

Note that $\sum_{k=1}^n \operatorname{tr}_{L/\mathbb{Q}}(\alpha_j\alpha_k)q_k = (T^TQ)_j = 0$, for any $j \in \{1, \ldots, n\}$. Hence, $U(\gamma, \beta) = 0$. As $\gamma \in L$ was taken to be arbitrary (and $\beta \in L \setminus \{0\}$), this contradicts the fact that U is non-degenerate. Hence, we have shown both directions of implication, and are done. □

Lemma 3.19. *Let $\alpha_1, \ldots, \alpha_n$ be elements of L. Then* $\operatorname{disc}(\alpha_1, \ldots, \alpha_n) = 0$ *if and only if $\alpha_1, \ldots, \alpha_n$ form a basis for L as a vector space over $\mathbb{Q}$.*

Proof. Suppose that the elements $\alpha_1, \ldots, \alpha_n$ are not a basis of L over $\mathbb{Q}$. Then they must be linearly dependent over $\mathbb{Q}$ (as there are $n = [L : \mathbb{Q}]$ of them). Hence, there exist some rationals $q_1, \ldots, q_n$ such that $q_j \neq 0$ for some $j \in \{1, \ldots, n\}$, and

$$\sum_{j=1}^n q_j\alpha_j = 0 \implies \sigma_i\left(\sum_{j=1}^n q_j\alpha_j\right) = \sum_{j=1}^n q_j\sigma_i(\alpha_j) = 0,$$

for any $i \in \{1, \ldots, n\}$. That is, the elements $\sigma_i(\alpha_1), \ldots, \sigma_i(\alpha_n)$ are linearly dependent over $\mathbb{C}$, for any $i \in \{1, \ldots, n\}$. But this just says that the columns of D are linearly dependent. Indeed, we have

$$q_1 \begin{bmatrix} \sigma_1(\alpha_1) \\ \vdots \\ \sigma_n(\alpha_1) \end{bmatrix} + \cdots + q_n \begin{bmatrix} \sigma_1(\alpha_n) \\ \vdots \\ \sigma_n(\alpha_n) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n q_j\sigma_1(\alpha_j) \\ \vdots \\ \sum_{j=1}^n q_j\sigma_n(\alpha_j) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

and $q_j \neq 0$ for at least one $j \in \{1, \ldots, n\}$. Hence, $\operatorname{disc}(\alpha_1, \ldots, \alpha_n) = \det(D)^2 = 0$. Now, suppose that the elements do form a basis of L over $\mathbb{Q}$. By Lemma 3.8, it suffices to show that the bilinear form U defined in said lemma is non-degenerate. Indeed, let $\beta \in L \setminus \{0\}$. Then

$$U(\beta^{-1}, \beta) = \operatorname{tr}_{L/\mathbb{Q}}(\beta^{-1}\beta) = \operatorname{tr}_{L/\mathbb{Q}}(1) = \sum_{k=1}^{n} \sigma_k(1) = n \neq 0.$$

Thus, U is non-degenerate (it is not possible for any non-zero element β to satisfy $U(\alpha, \beta) = 0$ for all $\alpha \in L$). □

Definition 3.20. An *integral basis* for $\mathcal{O}_L$ is a tuple $\alpha_1, \ldots, \alpha_n$ of elements of L such that

$$I = \bigoplus_{i=1}^{n} \mathbb{Z}\alpha_i.$$

Lemma 3.21. *Let $\alpha \in L$. Then there exists an integer $k \geq 1$ such that $k\alpha \in \mathcal{O}_L$.*

Proof. Firstly, there must exist a monic polynomial $f \in \mathbb{Q}[x]$ such that $f(\alpha) = 0$ (as $1, \alpha, \ldots, \alpha^n$ are linearly dependent over $\mathbb{Q}$). We can write

$$f(x) = b_n x^n + \cdots + b_1 x + b_0,$$

for some $b_0, \ldots, b_n \in \mathbb{Q}$, with $b_{\deg f} \neq 0$. Let $k = \operatorname{lcm}(b_0, \ldots, b_n)$ and let $g \in \mathbb{Q}[x]$ be defined by

$$g(x) = k^{\deg f} f(x/k).$$

Then, $g(k\alpha) = 0$ and g is monic. Thus, $k\alpha \in \mathcal{O}_L$. □

Lemma 3.22. *There exists an integral basis for $\mathcal{O}_L$.*

Proof. Let $\beta_1, \ldots, \beta_n$ be a basis of L over $\mathbb{Q}$. There exist integers $k_1, \ldots, k_n \geq 1$ such that $k_i\beta_i \in \mathcal{O}_L$ for each $i \in \{1, \ldots, n\}$. Thus $\gamma_1, \ldots, \gamma_n$ is a basis of L over $\mathbb{Q}$, where $\gamma_i = \operatorname{lcm}(k_1, \ldots, k_n)\beta_i$ for each $i \in \{1, \ldots, n\}$. Furthermore,

$$\left\{ \sum_{i=1}^{n} m_i \gamma_i : m_i \in \mathbb{Z} \right\} = \bigoplus_{i=1}^{n} \mathbb{Z}\gamma_i \subset \mathcal{O}_L,$$

as $\mathcal{O}_L$ is a ring. Now, in the previous lemma, we showed that the matrix T defined in Lemma 3.6 is invertible, as $\gamma_1, \ldots, \gamma_n$ is a basis for L over $\mathbb{Q}$. For each $i \in \{1, \ldots, n\}$, let γ_i^* be defined by

$$\gamma_i^* = \sum_{k=1}^{n} (T^T)^{-1}_{ki} \gamma_k.$$

Then,

$$\begin{aligned} U(\gamma_i^*, \gamma_j) = U\left(\sum_{k=1}^{n} (T^T)^{-1}_{ki} \gamma_k, \gamma_j \right) = \sum_{k=1}^{n} (T^T)^{-1}_{ki} U(\gamma_k, \gamma_j) &= \sum_{k=1}^{n} (T^T)^{-1}_{ki} \operatorname{tr}_{L/\mathbb{Q}}(\gamma_k \gamma_j) \\ &= \sum_{k=1}^{n} (T^T)^{-1}_{ki} T_{kj} = \delta_{ij}, \end{aligned}$$

as $\sum_{k=1}^n (T^T)^{-1}_{ki} T_{kj} = (((T^T)^{-1})^T T)_{ij} = (((T^{-1})^T)^T T)_{ij} = (T^{-1}T)_{ij} = (I_n)_{ij} = \delta_{ij}$. Assume that there exist rationals $q_1, \ldots, q_n \in \mathbb{Q}$ such that $\sum_{i=1}^n q_i \gamma_i^* = 0$. Then

$$0 = \sum_{k=1}^n \sigma_k(0) = \mathrm{tr}_{L/\mathbb{Q}}(0) = U(0, \gamma_j) = U\left(\sum_{i=1}^n q_i \gamma_i^*, \gamma_j\right) = \sum_{i=1}^n q_i U(\gamma_i^*, \gamma_j) = \sum_{i=1}^n q_i \delta_{ij} \\ = q_j,$$

for any $j \in \{1, \ldots, n\}$. Hence, $\gamma_1^*, \ldots, \gamma_n^*$ are linearly independent and thus form a basis of L over $\mathbb{Q}$. Now, let $\alpha \in \mathcal{O}_L$ be given. We can write $\alpha = \sum_{i=1}^n p_i \gamma_i^*$ for some rationals $p_1, \ldots, p_n$. Thus

$$\mathrm{tr}_{L/\mathbb{Q}}(\alpha\gamma_j) = U(\alpha, \gamma_j) = U\left(\sum_{i=1}^n p_i \gamma_i^*, \gamma_j\right) = \sum_{i=1}^n p_i U(\gamma_i^*, \gamma_j) = \sum_{i=1}^n p_i \delta_{ij} = p_j,$$

for each $j \in \{1, \ldots, n\}$. Recall that $\alpha, \gamma_j \in \mathcal{O}_L$. Hence, as $\mathcal{O}_L$ is a ring, $\alpha\gamma_j \in \mathcal{O}_L$. Thus, $\mathrm{tr}_{L/\mathbb{Q}}(\alpha\gamma_j) \in \mathbb{Z}$. Hence, $p_j \in \mathbb{Z}$ for all $j \in \{1, \ldots, n\}$, and

$$\mathcal{O}_L \subset \bigoplus_{i=1}^n \mathbb{Z}\gamma_i^*.$$

By the sandwich lemma, $\mathcal{O}_L \cong \mathbb{Z}^n$, and there exists an integral basis for $\mathcal{O}_L$. □

Definition 3.23. We define $\mathrm{disc}(\mathcal{O}_L) = \mathrm{disc}(\alpha_1, \ldots, \alpha_n)$, where $\alpha_1, \ldots, \alpha_n$ is any integral basis for $\mathcal{O}_L$.

Lemma 3.24. $\mathrm{disc}(\mathcal{O}_L) \neq 0$.

Proof. In the proof of Lemma 3.11, we constructed an integral basis for $\mathcal{O}_L$ that also formed a basis of L over $\mathbb{Q}$ (namely, $\gamma_1^*, \ldots, \gamma_n^*$). Hence, the desired result follows by Lemma 3.9. □

Now, let $\sigma_1 \ldots, \sigma_r$ denote the r real embeddings of L, and $\tau_1, \overline{\tau}_1, \ldots, \tau_s, \overline{\tau}_s$ denote the s conjugate pairs of complex embeddings. Identifying $\mathbb{C}$ with $\mathbb{R}^2$, we define a map $S : L \to \mathbb{R}^{r+2s} = \mathbb{R}^r \times \mathbb{C}^s$ by

$$S(\alpha) := (\sigma_1(\alpha), \ldots, \sigma_r(\alpha), \mathrm{Re}(\tau_1(\alpha)), \mathrm{Im}(\tau_1(\alpha)), \ldots, \mathrm{Re}(\tau_s(\alpha)), \mathrm{Im}(\tau_s(\alpha))).$$

Note that, as each embedding is a field homomorphism and by the respective additivity properties of Re and Im, S defines a group homomorphism. Moreover, S is injective, as clearly $S(\alpha) = 0 \iff \alpha = 0$.

Theorem 3.25. *$S(\mathcal{O}_L)$ is a lattice.*

Proof. Let $\alpha_1, \ldots, \alpha_n$ be an integral basis for $\mathcal{O}_L$. Then, as S is a homomorphism, we have

$$S(\mathcal{O}_L) = \bigoplus_{i=1}^n \mathbb{Z}S(\alpha_i).$$

Hence, showing that $S(\mathcal{O}_L)$ is a lattice amounts to showing that $S(\alpha_1), \ldots, S(\alpha_n)$ are linearly independent. Equivalently, we need to show that the matrix A whose jth column is defined by $S(\alpha_j)$ (for each $j \in \{1, \ldots, n\}$) has non-zero determinant. Note that

$$S(\alpha_j) = (\sigma_1(\alpha_j), \ldots, \sigma_r(\alpha_j), \mathrm{Re}(\tau_1(\alpha_j)), \mathrm{Im}(\tau_1(\alpha_j)), \ldots, \mathrm{Re}(\tau_s(\alpha_j)), \mathrm{Im}(\tau_s(\alpha_j))).$$

Let $B \in M_{2s}(\mathbb{C})$ be the block matrix

$$B = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & B_s \end{pmatrix},$$

where

$$B_j = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}.$$

for each $j \in \{1, \dots, n\}$. Let $M \in M_{r+2s}\mathbb{C}$ be the block matrix given by $M = \begin{pmatrix} I_r & 0 \\ 0 & B \end{pmatrix}$. Then, the jth column of MA is given by

$$(\sigma_1(\alpha_j), \dots, \sigma_r(\alpha_j), \tau_1(\alpha_j), \overline{\tau_1}(\alpha_j), \dots, \tau_s(\alpha_j), \overline{\tau_s}(\alpha_j)).$$

Furthermore,

$$\det(M)\det(A) = \det(MA) \implies (-2i)^s \det(A) = \det(MA).$$

But $\det(MA)^2 = \operatorname{disc}(\mathcal{O}_L)$. Thus,

$$(-2i)^{2s} \det(A)^2 = \operatorname{disc}(\mathcal{O}_L) \neq 0.$$

Hence, $\det(A) \neq 0$, and we are done. □

Lemma 3.26. $A(S(\mathcal{O}_L)) = \frac{1}{2^s}\sqrt{|\operatorname{disc}(\mathcal{O}_L)|}$.

Proof. Recall that

$$\begin{aligned} A(S(\mathcal{O}_L)) = \operatorname{vol}\left(\left\{\sum_{i=1}^n t_i S(\alpha_i) : t_i \in [0,1)\right\}\right) &= \operatorname{vol}\left(\left\{\sum_{i=1}^n t_i \sum_{j=1}^n A_{ji} e_j : t_i \in [0,1)\right\}\right) \\ &= \operatorname{vol}\left(\left\{\sum_{j=1}^n \left(\sum_{i=1}^n t_i A_{ji}\right) e_j : t_i \in [0,1)\right\}\right) \\ &= \operatorname{vol}\left(\{At : t \in [0,1)^n\}\right) = |\det(A)|, \end{aligned}$$

where A is the matrix defined in Theorem 3.3. Thus, the conclusion follows by the last calculation in the above lemma. □

Definition 3.27. Let $I \lhd \mathcal{O}_L$ be an ideal. The *norm* $N(I)$ of I is the index $[\mathcal{O}_L : I]$.

Lemma 3.28. *Let $I \lhd \mathcal{O}_L$ be a non-zero ideal. Then $N(I)$ is finite.*

Proof. First suppose that I is a principal ideal. We can write $I = (\beta)$ for some $\beta \in \mathcal{O}_L$. Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $\mathcal{O}_L$. Note then that

$$I = (\beta) = \{\gamma\beta : \gamma \in \mathcal{O}_L\} = \{\beta\gamma : \gamma \in \mathcal{O}_L\} = \beta\mathcal{O}_L = \bigoplus_{i=1}^n \mathbb{Z}\beta\alpha_i,$$

as $\mathcal{O}_L$ is a commutative ring. Hence, $I \cong \mathbb{Z}^n$. Now, suppose that I is any non-zero ideal. Let $\alpha \in I \setminus \{0\}$. Then we have the chain $(\alpha) \subset I \subset \mathcal{O}_L$ of inclusions of abelian groups. By the Sandwich Lemma, we have $I \cong \mathbb{Z}^n$, as $(\alpha) \lhd \mathcal{O}_L$ is principal. In either case, the Sandwich Lemma tells us that as $I \cong \mathbb{Z}^n \cong \mathcal{O}_L$, $\mathcal{O}_L/I$ is finite. □

Lemma 3.29. *Let $K \in \mathbb{N}$. Then there are only finitely many ideals $I \lhd \mathcal{O}_L$ such that $N(I) \leq K$.*

Proof. Let $N(I) = N$. By Lagrange's Theorem, we have that $N(\alpha + I) = I$ for any $\alpha \in \mathcal{O}_L$. Equivalently, $N\alpha \in I$ for any $\alpha \in \mathcal{O}_L$. Thus, $N \in I$ (as $1 \in \mathcal{O}_L$). Let $\pi : \mathcal{O}_L \to \mathcal{O}_L/(N)$ be the projection map $\alpha \mapsto \alpha + (N)$. Let $I \lhd \mathcal{O}_L$ be an ideal containing N. One can easily check that $\pi(I) \lhd \mathcal{O}_L/(N)$. Conversely, let $J \lhd \mathcal{O}_L/(N)$ be an ideal. Again, one can easily check that $\pi^{-1}(J) \lhd \mathcal{O}_L$. Furthermore, $N \in \pi^{-1}(J)$, as $N + (N) = 0 \in J$. Moreover,

$$\pi^{-1}(\pi(I)) = \pi^{-1}(\{i + (N) : i \in I\}) = I,$$

and

$$\pi(\pi^{-1}(J)) = \pi(\{\alpha \in \mathcal{O}_L : \alpha + (N) \in J\}) = J.$$

Thus, π gives a bijection between ideals of $\mathcal{O}_L$ containing N and ideals of $\mathcal{O}_L/(N)$. By the Sandwich Lemma, $\mathcal{O}_L/(N)$ is finite, and thus there are only finitely many ideals of $\mathcal{O}_L/(N)$. The one-to-one correspondence we have established tells us that there are only finitely many ideals I such that $N(I) = N$. We can easily generalise this statement to give the desired result. □

Note that the argument in Lemma 3.14 also shows that any non-zero ideal $I \lhd \mathcal{O}_L$ admits an integral basis. That is, we have

$$I = \bigoplus_{i=1}^{n} \mathbb{Z}\gamma_i$$

for some $\gamma_1, \ldots, \gamma_n \in I$. Hence, for a non-zero ideal $I \lhd \mathcal{O}_L$, we define

$$\operatorname{disc}(I) = \operatorname{disc}(\gamma_1, \ldots, \gamma_n),$$

where $\gamma_1, \ldots, \gamma_n \in I$ is any integral basis for I. Note that the discriminant is independent on the choice of integral basis, by the same reasoning as for $\mathcal{O}_L$.

Lemma 3.30. *If $I \lhd \mathcal{O}_L$ is a non-zero ideal, then* $\operatorname{disc}(I) = \operatorname{disc}(\mathcal{O}_L)N(I)^2$.

Proof. Let $\alpha_1, \ldots, \alpha_n$ be an integral basis for $\mathcal{O}_L$, and let $\gamma_1, \ldots, \gamma_n$ be an integral basis for I. For each $j \in \{1, \ldots, n\}$, we can write $\gamma_j = \sum_{k=1}^{n} B_{kj}\alpha_k$, for some $B_{1j}, \ldots, B_{nj} \in \mathbb{Z}$. Let $\tilde{B} \in M_n(\mathbb{Z})$ be the matrix defined by $\tilde{B}_{kj} = B_{kj}$. Let $A, C \in M_n(\mathbb{Z})$ be the matrices defined by $A_{ij} = \operatorname{tr}_{L/\mathbb{Q}}(\alpha_i\alpha_j)$ and $C_{ij} = \operatorname{tr}_{L/\mathbb{Q}}(\gamma_i\gamma_j)$, respectively. Then,

$$\begin{aligned} C_{ij} = \operatorname{tr}_{L/\mathbb{Q}}\left(\sum_{k=1}^{n} B_{ki}\alpha_k \sum_{\ell=1}^{n} B_{\ell j}\alpha_\ell\right) &= \sum_{k=1}^{n}\sum_{\ell=1}^{n} B_{ki}B_{\ell j}\operatorname{tr}_{L/\mathbb{Q}}(\alpha_k, \alpha_\ell) \\ &= (B^T AB)_{ij}. \end{aligned}$$

Thus, $C = B^T AB$ and hence $\det C = \det(A)\det(B)^2 = \det(B)^2 \operatorname{disc}(\mathcal{O}_L)$. Since $B \in M_n(\mathbb{Z})$, and $\mathbb{Z}$ is a Euclidean domain, we can put the matrix B into Smith normal form. That is, we can perform elementary row and column operations to B and obtain the matrix $\operatorname{diag}(d_1, \ldots, d_n)$, where $d_i \mid d_{i+1}$ for all $i \in \{1, \ldots, k-1\}$. Thus, we have that $\mathbb{Z}^n/B\mathbb{Z}^n \cong \bigoplus_{i=1}^{n} \mathbb{Z}/d_i\mathbb{Z}$. Now, we have the projection map $\pi : \mathcal{O}_L \to \mathcal{O}_L/I$. We can identify $\mathcal{O}_L$ with $\mathbb{Z}^n$ via the map

$(c_1, \dots, c_n) \mapsto \sum_{i=1}^n c_i\alpha_i$. This gives us a surjective homomorphism $\tilde{\pi} : \mathbb{Z}^n \to \mathcal{O}_L/I$. Moreover, for any $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$,we then have

$$\begin{aligned} x \in \ker\tilde{\pi} \iff \sum_{i=1}^n x_i\alpha_i \in I &\iff \sum_{i=1}^n x_i\alpha_i = \sum_{j=1}^n y_j\gamma_j \text{ for some } y_1, \dots, y_n \in \mathbb{Z} \\ &\iff \sum_{i=1}^n x_i\alpha_i = \sum_{j=1}^n y_j \sum_{k=1}^n B_{kj}\alpha_k \\ &\iff \sum_{i=1}^n x_i\alpha_i = \sum_{k=1}^n \left(\sum_{j=1}^n B_{kj}y_j\right)\alpha_k. \end{aligned}$$

Comparing coefficients, we see that $x \in B\mathbb{Z}^n$. Thus, we have $\mathbb{Z}^n/B\mathbb{Z}^n \cong \mathcal{O}_L/I$ by the first isomorphism theorem. Thus $\bigoplus_{i=1}^n \mathbb{Z}/d_i\mathbb{Z} \cong \mathbb{Z}^n/B\mathbb{Z}^n \cong \mathcal{O}_L/I$. Recall that $\mathcal{O}_L/I$ is finite (and also abelian). Hence, by the Fundamental Theorem of Finite Abelian Groups, we must have that $N(I) = [\mathcal{O}_L : I] = |\mathcal{O}_L/I| = \prod_{i=1}^n d_i = |\det B|$. The desired result follows by our previous working. □

Corollary 3.31. *Let $\beta \in \mathcal{O}_L$ be non-zero, and let $I = (\beta)$. Then $N(I) = |N_{L/\mathbb{Q}}(\beta)|$.*

Proof. Let $\alpha_1, \dots, \alpha_n$ be an integral basis for $\mathcal{O}_L$. The argument of Lemma 3.14 shows that $\beta\alpha_1, \dots \beta\alpha_n$ is an integral basis for I. Let $\sigma_1, \dots, \sigma_n$ denote the n complex embeddings of L. Then, $\operatorname{disc}(I) = \det(D)^2$, where $D_{ij} = \sigma_i(\beta\alpha_j)$. Note that $D_{ij} = \sigma_i(\beta)\sigma_i(\alpha_j)$. Hence,

$$\operatorname{disc}(I) = \det(D)^2 = \left(\prod_{i=1}^n \sigma_i(\beta)\right)^2 \det\left(\tilde{D}\right)^2 = N_{L/\mathbb{Q}}(\beta)^2 \operatorname{disc}(\mathcal{O}_L),$$

where $\tilde{D}_{ij} = \sigma_i(\alpha_j)$. By Lemma 3.16, we then have $N_{L/\mathbb{Q}}(\beta)^2 = N(I)^2$, and the desired conclusion follows. □

4. Proof of Dirichlet's Unit Theorem

Let L be a number field. Let $\sigma_1, \dots, \sigma_r, \tau_1, \overline{\tau_1}, \dots, \tau_s, \overline{\tau_s}$ denote the complex embeddings of L. Consider the map $\ell : L^\times \to \mathbb{R}^{r+s}$, defined by

$$\ell(\alpha) = (\log|\sigma_1(\alpha)|, \dots, \log|\sigma_r(\alpha)|, 2\log|\tau_1(\alpha)|, \dots, 2\log|\tau_s(\alpha)|).$$

Note firstly that ℓ is well-defined, as the field homomorphisms $\sigma_1, \dots, \sigma_r, \tau_1, \overline{\tau_1}, \dots, \tau_s, \overline{\tau_s}$ are injective, and $0 \notin L^\times$ (thus we can take the inner logs without any issues). Furthermore, the additivity property of each of the embeddings and the natural logarithm mean that ℓ defines a group homomorphism. Now, we claim that

$$\ell(\mathcal{O}_L^\times) \subseteq H = \left\{ (x_1, \dots, x_{r+s}) : \sum_{i=1}^{r+s} x_i = 0 \right\}$$

Indeed, let $\alpha \in \mathcal{O}_L^\times$. We have that

$$\begin{aligned} N_{L/\mathbb{Q}}(\alpha) &= \prod_{i=1}^{r} \sigma_i(\alpha) \prod_{i=1}^{s} \tau_i(\alpha)\overline{\tau_i}(\alpha) \\ &= \prod_{i=1}^{r} \sigma_i(\alpha) \prod_{i=1}^{s} |\tau_i(\alpha)|^2 = 1. \end{aligned}$$

Taking the log of the absolute values, we have

$$\begin{aligned} \log|N_{L/\mathbb{Q}}(\alpha)| = \sum_{i=1}^{r} \log|\sigma_i(\alpha)| + 2\sum_{i=1}^{s} \log|\tau_i(\alpha)| = 0 \\ \implies \ell(\alpha) \in H. \end{aligned}$$

Thus, $\ell(\mathcal{O}_L^\times) \subseteq H$. Now, note that H is a subspace of $\mathbb{R}^{r+s}$ (the summation property of the elements of H is invariant under scalar multiplication, and H is clearly closed under addition). Note also that $H = \ker \Sigma$, where $\Sigma : \mathbb{R}^{r+s} \to \mathbb{R}$ is defined by

$$\Sigma(x_1, \dots, x_{r+s}) = \sum_{i=1}^{r+s} x_i.$$

Furthermore, it is clear that Σ is a linear mapping. Note also that Σ is surjective ($\Sigma(t, 0, \dots, 0) = t$ for any $t \in \mathbb{R}$), and thus $\dim \operatorname{im} \Sigma = \dim \mathbb{R} = 1$. Hence, by the rank-nullity theorem, we have that

$$\dim H = \dim \ker \Sigma = \dim(\mathbb{R}^{r+s}) - \dim(\operatorname{im} \Sigma) = r + s - 1.$$

Thus, H is a subspace of $\mathbb{R}^{r+s}$ of dimension $r + s - 1$.

Lemma 4.1. *Let $1 \le k \le r + s$ be an integer, and $\alpha \in \mathcal{O}_L \setminus \{0\}$ be given. Let $\ell(\alpha) = (a_1, \dots, a_{r+s})$. Then there exists $\beta \in \mathcal{O}_L \setminus \{0\}$ such that:*

(1) $N_{L/\mathbb{Q}}(\beta) \le \left(\frac{2}{\pi}\right)^s \sqrt{|\operatorname{disc}(\mathcal{O}_L)|}$;

(2) Let $\ell(\beta) = (b_1, \dots, b_{r+s})$. Then $b_i < a_i$ if $i \neq k$.

Proof. Let $E \subset \mathbb{R}^n = \mathbb{R}^r \times \mathbb{C}^s$ be the region defined by

$$|y_1| \le c_1, \dots, |y_r| \le c_r, |z_1|^2 \le c_{r+1}, \dots, |z_s|^2 \le c_{r+s}.$$

where $c_i \in \mathbb{R}^+$ are the positive real numbers defined by

$$0 < c_i < e^{a_i} \quad (i \neq k),$$

and

$$\prod_{i=1}^{r+s} c_i = \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_L)|}.$$

Firstly, it is important to note that we *can* choose such real numbers. Indeed, we can define

$$c_k := \frac{\left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_L)|}}{\prod_{i \neq k} c_i},$$

and the conditions are satisfied (as the above expression is well-defined, since the denominator is non-zero). Now, $\mathrm{vol}(\partial E) = 0$. Furthermore, E is closed and bounded, and hence compact (by the Heine-Borel theorem). Hence, as $S(\mathcal{O}_L)$ is a lattice, and

$$\begin{aligned}\mathrm{vol}(E) = \prod_{i=1}^{r} 2c_i \prod_{i=1}^{s} \pi c_{r+i} = 2^r \pi^s \prod_{i=1}^{r+s} c_i &= 2^r \pi^s \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_L)|} \\ &= 2^{r+s} \sqrt{|\mathrm{disc}(\mathcal{O}_L)|} \\ &= 2^{r+s}(2^s A(S(\mathcal{O}_L)) \\ &= 2^{r+2s} A(S(\mathcal{O}_L)).\end{aligned}$$

Thus, $\mathrm{vol}(E) = 2^n A(S(\mathcal{O}_L))$, and thus (by Minkowski's Theorem), there exists some $\beta \in \mathcal{O}_L \setminus \{0\}$ such that $S(\beta) \in E$. Thus,

$$N_{L/\mathbb{Q}}(\beta) = \prod_{i=1}^{r} \sigma_i(\beta) \prod_{i=1}^{s} |\tau_i(\beta)|^2 \leq \prod_{i=1}^{r+s} c_i = \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_L)|},$$

and condition (1) is satisfied. Furthermore,

$$b_i = \begin{cases} \log|\sigma_i(\beta)| \leq \log|c_i| \leq \log|e^{a_i}| = a_i & \text{if } r \geq i \neq k, \\ \log|\tau_i(\beta)|^2 \leq \log|c_i| \leq \log|e^{a_i}| = a_i & \text{if } s \leq i \neq k. \end{cases}$$

This gives condition (2). □

Corollary 4.2. *Let $1 \leq k \leq r+s$ be an integer. Then there exists an element $\varepsilon \in \mathcal{O}_L^\times$ such that, writing $\ell(\varepsilon) = (e_1, \ldots, e_{r+s})$, we have $e_i > 0$ if $i \neq k$ and $e_k < 0$.*

Proof. Take an arbitrary element $\alpha \in \mathcal{O}_L \setminus \{0\}$. We can apply Lemma 4.1 to obtain an element $\alpha_1 \in \mathcal{O}_L \setminus \{0\}$ that satisfies the conditions of said lemma. We can then apply Lemma 4.1 to α_1, and so on, and we can therefore obtain an infinite set $\{\alpha_j\}_{j \in \mathbb{N}}$ of non-zero elements of $\mathcal{O}_L$ such that

(1) $N_{L/\mathbb{Q}}(\alpha_j) \leq \left(\frac{2}{\pi}\right)^s \sqrt{|\mathrm{disc}(\mathcal{O}_L)|}$.

(2) Let $\ell(\alpha_{j+1}) = (b_1, \ldots, b_{r+s})$. Then $b_i < a_i$ if $i \neq k$, where $\ell(\alpha_j) = (a_1, \ldots, a_{r+s})$

for any $j \in \mathbb{N}$. Since the set $\{\alpha_j\}_{j \in \mathbb{N}}$ is infinite, and each of the elements is bounded in norm, we must have (by the pigeonhole principle) that $(\alpha_j) = (a_{j'})$ for some $j < j'$, as there are only finitely many ideals of order $N_{L/\mathbb{Q}}(\alpha_j) = N((\alpha_j))$. Thus, we have $\alpha_j = x\alpha_{j'}$ and $\alpha_{j'} = y\alpha_j$ for some $x, y \in \mathcal{O}_L$. Hence, $\alpha_j\alpha_{j'}^{-1} = x$. Furthermore, we have

$$\alpha_j = x(y\alpha_j) = (xy)\alpha_j \implies xy = 1 \implies \alpha_j\alpha_{j'}^{-1} = x \in \mathcal{O}_L^\times.$$

Moreover, we can write $j' = j + m$ for some $m \in \mathbb{N}$, and hence

$$\ell(\alpha_j\alpha_{j'}^{-1}) = \sum_{i=0}^{m-1} \ell(\alpha_{(j+i)}(\alpha_{j+(i+1)})^{-1}) = \sum_{i=0}^{m-1} \left(\ell(\alpha_{(j+i)}) - \ell(\alpha_{j+(i+1)})\right).$$

Thus, (2) implies that each summand is a vector satisfying (2), and hence the summation itself is also. Write $\varepsilon = \alpha_j\alpha_{j'}^{-1}$. Then, as $\varepsilon \in \mathcal{O}_L^\times$, we have

$$\begin{aligned} |N_{L/\mathbb{Q}}(\varepsilon)| = 1 &\implies \left|\prod_{i=1}^{r} \sigma_i(\varepsilon) \prod_{i=1}^{s} |\tau_i(\varepsilon)|^2\right| = \prod_{i=1}^{r} |\sigma_i(\varepsilon)| \prod_{i=1}^{s} |\tau_i(\varepsilon)|^2 = 1 \\ &\implies \log\left(\prod_{i=1}^{r} |\sigma_i(\varepsilon)| \prod_{i=1}^{s} |\tau_i(\varepsilon)|^2\right) = \underbrace{\sum_{i=1}^{r} \log|\sigma_i(\varepsilon)| + \sum_{i=1}^{s} 2\log|\tau_i(\varepsilon)|}_{=\sum_{i=1}^{r+s} e_i} = 0. \end{aligned}$$

Note that, as $e_i \neq 0$ for $i \neq k$, the above forces $e_k = -\sum_{i \neq k} e_k < 0$. Thus, ε is such an element, and we are done. □

Lemma 4.3. *Let $N \geq 1$ be an integer and let $A \in M_N(\mathbb{R})$ be a matrix satisfying the following conditions:*

(1) For each $j \in \{1, \dots, N\}$, $\sum_{i=1}^{N} A_{ij} = 0$.

(2) For all $i, j \in \{1, \dots N\}$, we have $A_{ij} > 0$ if $i = j$ and $A_{ij} < 0$ if $i \neq j$.

Then A has rank $N - 1$.

Proof. Let $T : \mathbb{R}^N \to \mathbb{R}^N$ be the linear mapping defined by $Tx = Ax$. Consider the vector $\mathbf{1} = (1, \dots, 1)^T \in \mathbb{R}^N$. We have that

$$(A\mathbf{1})_i = \sum_{j=1}^{N} A_{ij} = 0$$

for each $i \in \{1, \dots N\}$. Hence, $\mathbf{1} \in \ker T$ and $\dim(\ker T) \geq 1$. By the rank-nullity theorem, we have that $\dim(\operatorname{im} T) \leq N - 1$. Thus A has rank at most $N - 1$. Now, assume that there exist some real numbers $t_1, \dots, t_{N-1}$ such that $t_i \neq 0$ for at least one $i \in \{1, \dots, N-1\}$ and that $\sum_{i=1}^{N-1} t_i A_{ij} = 0$ for each $j \in \{1, \dots N\}$. Now, since at least one of the t_is is non-zero, we can divide each t_i by $\max\{t_i \neq 0 : i \in \{1, \dots, N-1\}\}$. Hence, there exists $k \in \{1, \dots, N-1\}$ such that $t_k = 1$, and $t_i \leq 1$ for all $i \in \{1, \dots N-1\}$. Then we have that

$$0 = \sum_{i=1}^{N-1} t_i A_{ik} \geq \sum_{i=1}^{N-1} A_{ik} > \sum_{i=1}^{N} A_{ik} = 0,$$

a contradiction. Note that both inequalities above follow by (2), as we have that $A_{Nk} < 0$ and $k \in \{1, \dots N-1\}$ (so $k \neq N$). Thus, no such real numbers exist, and thus the first $N - 1$ rows of A are linearly independent. It follows that the rank of A is $N - 1$. □

Lemma 4.4. *Let $B > 0$ be a real number, and let*

$$X_B := \{\alpha \in \mathcal{O}_L : \textit{for all embeddings } \sigma : L \to \mathbb{C}, |\sigma(\alpha)| \leq B\}.$$

Then X_B is finite.

Proof. Note that

$$S(X_B) = S(\mathcal{O}_L) \cap [-B, B]^r \times \{z \in \mathbb{C}^s : |z_j| \leq B \text{ for each } j \in \{1, \dots s\}\},$$

and the set on the right-hand side is compact. Thus, as $S(\mathcal{O}_L)$ is a lattice, it follows that $S(X_B)$ is finite. Since S is injective, it follows that X_B is finite. □

Proposition 4.5. *$\ell(\mathcal{O}_L^\times)$ is a lattice in H.*

Proof. Recall firstly that the image of $\mathcal{O}_L^\times$ under ℓ must be a subgroup of $\mathbb{R}^{r+s}$, as the image of a subgroup under a group homomorphism is a subgroup. Furthermore, $\mathbb{R}^{r+s}$ is abelian, and thus all subgroups of $\mathbb{R}^{r+s}$ are abelian.

We first show that $\ell(\mathcal{O}_L^\times)$ spans H. Corollary 4.1 implies the existence of vectors $v_1, \dots, v_{r+s} \in \ell(\mathcal{O}_L^\times)$ such that the ith entry of v_j is strictly positive if $i \neq j$, and negative otherwise, for each $j \in \{1, \dots, r+s\}$. Let $A \in M_{r+s}(\mathbb{R})$ be the matrix with column j given by v_j. Then A satisfies the conditions of Lemma 4.2, and its rank is hence $r + s - 1$. Earlier, we computed that $\dim H = r + s - 1$, and thus $\ell(\mathcal{O}_L^\times)$ spans H.

Recall that any spanning set of a finite-dimensional vector space contains a basis. Thus, we may choose vectors $v_1, \dots, v_{r+s-1} \in \ell(\mathcal{O}_L^\times)$ that form a basis of H as a vector space over $\mathbb{R}$. Define

$$\Lambda = \bigoplus_{i=1}^{r+s-1} \mathbb{Z} v_i.$$

Then $v_1, \dots, v_{r+s-1}$ span Λ, and $\Lambda \subset \ell(\mathcal{O}_L^\times)$. Let $P \subset H$ be defined by

$$P = \left\{ \sum_{i=1}^{r+s-1} t_i v_i : t_i \in [0, 1] \text{ for each } i \in \{1, \dots, r+s-1\} \right\}.$$

Note that P is indeed a subset of H, as we have established that the vectors $v_1, \dots, v_{r+s-1}$ form a basis of H as a vector space over $\mathbb{R}$. Now, suppose that $\ell(\alpha) \in P$ for some $\alpha \in \mathcal{O}_L^\times$. Then, by the definition of ℓ, $|\sigma(\alpha)|$ is bounded for any embedding $\sigma : L \to \mathbb{C}$, and furthermore this bound is independent of the embedding (as P is bounded). By Lemma 4.3, we can conclude that $P \cap \ell(\mathcal{O}_L^\times)$ is finite. By a similar argument to that of the proof of Minkowski's Theorem, we can write $x = \lambda + p$ for some $\lambda \in \Lambda$ and $p \in P$, for each $x \in \ell(\mathcal{O}_L^\times)$. Note then that

$$p = (\lambda - x) \in \ell(\mathcal{O}_L^\times) \cap P,$$

as $\Lambda \subset \ell(\mathcal{O}_L^\times)$, and $\ell(\mathcal{O}_L^\times)$ is a group. Hence, for each $x \in \ell(\mathcal{O}_L^\times)$, we can write $x = \lambda + p$ for some $p \in \ell(\mathcal{O}_L^\times) \cap P$. As $\ell(\mathcal{O}_L^\times) \cap P$ is finite, this means that the number of cosets of $\Lambda \subset \ell(\mathcal{O}_L^\times)$ is finite. Note that we can consider the quotient group $\ell(\mathcal{O}_L)/\Lambda$, as $\ell(\mathcal{O}_L^\times)$ is abelian, so all of its subgroups are normal. Hence, we can write $[\ell(\mathcal{O}_L^\times) : \Lambda] = N$ for some $N \in \mathbb{N}$. By Lagrange's

Theorem, we then have that $N(x+\Lambda) = 0_{\ell(\mathcal{O}_L^\times)}$, for any $x \in \ell(\mathcal{O}_L^\times)$. This is equivalent to stating that $N\ell(\mathcal{O}_L^\times) \subset \Lambda$. Hence, we have that

$$\Lambda \subset \ell(\mathcal{O}_L^\times) \subset \frac{1}{N}\Lambda.$$

Note that scaling a lattice by a non-zero real number does not affect its structure (this can be made precise using our previous arguments with homeomorphisms). In other words, if we scale a lattice, then we still end up with a lattice. Hence, we can apply the Sandwich Lemma to $\ell(\mathcal{O}_L^\times)$, and conclude that $\ell(\mathcal{O}_L^\times) \cong \mathbb{Z}^{r+s-1}$. It follows that $\ell(\mathcal{O}_L^\times)$ is a lattice in H. □

Theorem 4.6 (Dirichlet's Unit Theorem).

The group μ_L is finite and cyclic, and we have the isomorphism

$$\mathcal{O}_L^\times \cong \mu_L \times \mathbb{Z}^{r+s-1}.$$

Proof. Note firstly that $\ker \ell \subset X_1$, so it is finite. As $\ker \ell$ is a subgroup, then this means that all of its elements are of finite order. Hence, $\ker \ell \subseteq \mu_L$. Conversely, if $\alpha \in \mu_L$, then $\alpha^N = 1$ for some $N \in \mathbb{N}$, and (as ℓ is a group homomorphism, so the additive identity 1 of L is mapped to 0)

$$\ell(\alpha^N) = N\ell(\alpha) = 0 \implies \ell(\alpha) = 0.$$

It follows that $\mu_L \subseteq \ker \ell$ and hence that $\mu_L = \ker \ell$. Thus, μ_L is finite and hence cyclic, as any finite subgroup of the group of roots of unity in $\mathbb{C}$ is cyclic. Now, let $u_1, \ldots, u_{r+s-1} \in \mathcal{O}_L^\times$ be elements whose image under ℓ forms a $\mathbb{Z}$-basis of $\ell(\mathcal{O}_L^\times)$. Let $f : \mu_L \times \mathbb{Z}^{r+s-1} \to \mathcal{O}_L^\times$ be the mapping defined by

$$f(w, a_1, \ldots, a_{r+s-1}) = wu_1^{a_1} \cdots u_{r+s-1}^{a_{r+s-1}}.$$

Firstly,

$$\begin{aligned} f((w_1, a_1, \ldots, a_{r+s-1}) \cdot (w_2, b_1, \ldots, b_{r+s-1})) &= (w_1 w_2 u_1^{a_1+b_1} \cdots u_{r+s-1}^{a_{r+s-1}+b_{r+s-1}}) \\ &= (wu_1^{a_1} \cdots u_{r+s-1}^{a_{r+s-1}})(w_2 u_1^{b_1} \cdots u_{r+s-1}^{b_{r+s-1}}) \\ &= f(w_1, a_1, \ldots, a_{r+s-1}) f(w_2, b_1, \ldots, b_{r+s-1})), \end{aligned}$$

so f is a homomorphism. Note that the second step above uses the fact that $\mathcal{O}_L^\times$ is abelian. Furthermore, suppose $(w, a_1, \ldots, a_{r+s-1}) \in \ker f$. Then, we have that

$$0 = \ell(1) = \ell(wu_1^{a_1} \cdots u_{r+s-1}^{a_{r+s-1}}) = \ell(w) + \sum_{i=1}^{r+s-1} \ell(u_i^{a_i}) = \sum_{i=1}^{r+s-1} a_i \ell(u_i).$$

Since $u_1, \ldots, u_{r+s-1} \in \mathcal{O}_L^\times$ are elements whose image under ℓ forms a $\mathbb{Z}$-basis of $\ell(\mathcal{O}_L^\times)$, we must have $a_i = 0$ for all $i \in \{1, \ldots, r+s-1\}$. Hence, $\ker f = (1, 0, \ldots, 0)$. Thus, only element of $\ker f$ is the identity, and f is thus injective. Finally, let $\alpha \in \mathcal{O}_L^\times$ be given. Then $\ell(\alpha) = \sum_{i=1}^{r+s-1} b_i \ell(u_i)$, for some $b_1, \ldots, b_{r+s-1} \in \mathbb{Z}$. Let $w = \alpha \prod_{i=1}^{r+s-1} u_i^{-b_i}$. Then

$$\ell(w) = \ell(\alpha) - \sum_{i=1}^{r+s-1} b_i \ell(u_i) = 0,$$

and $w \in \ker \ell = \mu_L$. Note then that

$$f(w, b_1, \ldots, b_{r+s-1}) = wu_1^{b_1} \cdots u_{r+s-1}^{b_{r+s-1}} = \alpha,$$

and f is thus surjective. Hence, f is an isomorphism, and the proof of Dirichlet's Unit Theorem is complete. $\square$